UNIVERSITÉ DE LIÈGE
Faculté des sciences
Sciences géographiques



Prototypage d'un système d'information géographique urbain pour la gestion des risques naturels

Bamenda-ville (Cameroun)

Mémoire présenté par
Daniel DEMONCEAU
pour l'obtention du diplôme de
master en sciences géographiques
option géomatique et géométrologie

Année académique
2009-2010

UNIVERSITÉ DE LIÈGE
Faculté des sciences
Sciences géographiques



Prototypage d'un système d'information géographique urbain pour la gestion des risques naturels

Bamenda-ville (Cameroun)

Mémoire présenté par
Daniel DEMONCEAU
pour l'obtention du diplôme de
master en sciences géographiques
option géomatique et géométrologie

Année académique
2009-2010

Je tiens tout particulièrement à dédier ce travail à ma grand-mère paternelle qui est décédée pendant la rédaction du mémoire. Elle fut toujours là pour nous, surtout dans les temps difficiles, et je regrette qu'elle n'ait pas vécu l'aboutissement de ce travail. Merci à elle.

Mes plus vifs remerciements vont à mon promoteur, Monsieur Jean-Paul DONNAY, pour avoir accepté de superviser ce mémoire ainsi que pour le temps qu'il a consacré à me prodiguer ses nombreux conseils.

Je voudrais remercier Monsieur Roland BILLEN et Monsieur François PETIT, lecteurs de ce mémoire, d'avoir accepté de faire partie du jury.

Je remercie également Monsieur Emmanuel Mba NYAMBOD, étudiant camerounais, ayant sollicité notre aide et avec qui nous avons mené à bien ce projet de collaboration qui n'aurait jamais vu le jour sans lui.

J'exprime ma reconnaissance à l'Université de Liège ainsi qu'à la Vrije Universiteit Brussel qui ont supporté les frais d'acquisition des données nécessaires à la réalisation de ce travail.

Je voudrais exprimer ma gratitude à Monsieur Marc BINARD, Monsieur Jean-PauL KASPRZYK et Monsieur Yves CORNET ainsi qu'à l'ensemble du personnel du laboratoire SURFACES pour l'aide qu'ils m'ont apportée lors de la réalisation des aspects plus techniques du travail.

J'adresse mes chaleureux remerciements à Martine, Christian et Jenny pour avoir effectué la relecture de l'entièreté du mémoire.

Enfin, je remercie ma famille et principalement Sylvie ainsi que toutes les personnes qui m'ont soutenu, de près ou de loin, tout au long de ce travail.

Table des matières

1. Introduction ….. 4
1.1. Introduction générale ….. 4
1.2. Intérêt et qualité d'un SIG urbain dans les pays en voie de développement …...................................... 5
1.3. Présentation du travail ….. 6
2. Objectifs de la recherche 9
2.1. Un travail en collaboration 9
2.1.1. Présentation générale ….. 9
2.1.2. Mise en contexte 10
2.1.3. Conception du SIG 12
2.2 Conception du SIG sur le territoire considéré selon standards et couts minimaux …........................... 14
2.2.1. Transformation des données existantes ….. 14
2.2.2. Modèle conceptuel de données ….. 15
2.2.3. Implémentation d'un prototype en *OpenAccess* ….. 16
2.3. Conception du Projet 17
3. Contextualisation 19
3.1. Démarche externe ….. 19
3.2. Sources de données 20
3.3. Hypothèse de travail 23
4. Analyse des données 24
4.1. Introduction 24
4.2. Analyse de la qualité des données brutes 25
4.2.1. Image aérienne 25
4.2.2. Les données Shapefile 26
4.3. Rapport qualité 35
4.4. Sélection de données utiles 36
4.5. Construction des classes ….. 37
4.5.1. Paquetage « Administratif » 38
4.5.2. Paquetage « Affectation de sol » 39
4.5.3. Paquetage « Buildings » 40
4.5.4. Paquetage « Hydrographie » 44
4.5.5. Paquetage « Relief » ….. 44
4.5.6. Paquetage « Voiries » 45
4.6. Renommage 47
4.7. Catalogue des données 47
5. Modélisation des données 49
5.1. Modèle conceptuel de données 51
5.1.1. Associations et cardinalités 51
5.1.2. Vérification du modèle 52
5.2. Modèle logique de données ….. 53
5.2.1. Création des classes et des relations ….. 54
5.3. Complétion de la classe Buildings 58
6. Prototype avec data loading ….. 62
6.1. Modèle physique 62
6.1.1. Évaluation du matériel informatique ….. 62
6.1.2. Création de la base de données spatiale ….. 63
6.1.3. Implémentation dans le SGBD ….. 63
6.2. Validation des données opérationnelles ….. 68
6.3. Accès pour une application SIG 72
7. Exploitation du prototype ….. 73
7.1. Visualisation 73
7.2. Enrichissement de données 75
7.3. Exportation de données ….. 76
8. Conclusion 77
8.1. Évaluation ….. 77
8.2. Perspectives ….. 80
9. Bibliographie 81
10. Annexes 84

1 Introduction

1.1 Introduction générale

Les catastrophes naturelles menaçant les populations vivant en tissu urbain sont relativement fréquentes et souvent particulièrement dévastatrices.
Parmi les plus récentes et les plus spectaculaires, nous pouvons citer :

- les inondations causées par l'ouragan Katrina faisant 1.322 morts dans le sud des États-Unis en 2005 ;
- les fortes pluies de mousson provoquant la mort de 1.275 personnes en Inde en novembre 2008 ;
- les inondations provoquées par d'abondantes pluies torrentielles causant la mort de 123 personnes et causant d'importantes pertes matérielles dont la destruction de 10.785 maisons et de 10.850 véhicules à Jeddah, en Arabie Saoudite, en novembre 2009 ;
- les tremblements de terre :
  - de Bam, en Iran, où l'on dénombra 26.200 morts en décembre 2003 ;
  - de Sumatra, en Indonésie, à l'origine du tsunami dévastateur qui fit 222.000 morts en décembre 2004 ;
  - de Port-au-Prince, en Haïti, où l'on ne compta pas moins de 230.000 morts ainsi que de multiples estropiés et d'innombrables sans abris en janvier 2010 ;
  - de Conception au Chili, provoquant la mort de 497 personnes en février 2010 ;
- les glissements de terrain suite à l'éruption du volcan Nevado del Ruiz qui fit fondre la neige dévalant les pentes de son cône en direction de la ville d'Amero, en Colombie, qui firent 24.000 morts en novembre 1985 ;
- les glissements de terrain suite à des pluies torrentielles à Vargas, au Venezuela, qui provoquèrent la mort d'au moins 15.000 personnes en décembre 1999.

Ces quelques exemples montrent combien les villes restent vulnérables face aux catastrophes naturelles, que ce soit en pays industrialisés ou en pays en voie de développement.

Pour diminuer les risques, il serait bon de mettre l'accent sur la prévention.

A cet effet, une collecte adéquate et une exploitation correcte de multiples données géographiques s'avèrent nécessaires afin de pouvoir réaliser des planifications en matière d'aménagement dans la gestion territoriale. Selon PENNOBER (2005), les meilleurs outils à utiliser sont, dans ce cas, ceux de la cartographie car ils permettent le mieux d'analyser les besoins et facilitent ainsi la compréhension de la problématique.

En tissu urbain, lorsque la zone d'étude est de grande taille, la quantité d'informations à traiter peut devenir rapidement conséquente, et alors un système d'information géographique (SIG) s'avère être un excellent outil. Lorsque toutes les données requises sont disponibles, il facilite l'analyse du risque pour un territoire donné et en réduit considérablement le temps d'exécution. Sans lui, l'analyse des données et la production de cartes seraient bien trop lourdes et fastidieuses à gérer.

1.2 **Intérêt et qualité d'un SIG urbain dans les pays en voie de développement**

Afin de démontrer l'intérêt qu'ont les pays en voie de développement pour l'utilisation des nouvelles technologies dans la gestion territoriale, le Journal Statistique Africain a mené une enquête auprès des pays du continent noir (SANGA & DOSSO 2007). Il en ressort que la plupart de ces pays, soit disposent actuellement d'un SIG, soit en planifient l'introduction. Seule la Tanzanie ne prévoit pas de l'introduire. Cet article tend à promouvoir davantage l'utilisation de ce type d'outil dans ces contrées défavorisées afin de permettre d'étudier l'évolution des indices de pauvreté.

Un exemple d'application concrète des nouvelles technologies dans la gestion territoriale en Afrique est la conception d'un SIG dans le cadre de la planification de la gestion des côtes en Afrique de l'Ouest (PENNOBER *et al*. 2005). Elle découle du souhait de la Guinée-Bissau de préserver ses riches littoraux aux points de vue physique, biologique, et culturel. La conception de ce SIG s'inscrit dans un projet de développement durable des zones menacées par les pressions humaine, touristique et industrielle qui engendrent une déforestation massive, une réduction importante de la biosphère, une exploitation excessive des ressources halieutiques et l'apparition de « sites poubelle » tels que des décharges de produits toxiques. Ce SIG est constitué de deux bases de données, l'une pour la planification de la gestion des provinces côtières, l'autre pour la planification des missions de conservation de la biosphère dans l'archipel des îles Bissagos.

Créé en partenariat par de nombreux pays africains de l'Ouest et du Sud du continent principalement, le réseau SIGAFRIQUE a pour but de « *contribuer à l'effort de réduction de la pauvreté en développant des politiques régionales africaines basées sur une valorisation et une diffusion de l'information en Sciences de la Terre du continent africain et destinées, entre autres, à la promotion de l'ensemble des ressources minérales, hydrogéologiques et au renforcement des compétences des services géologiques nationau*x » (SIGAFRIQUE 2007). Bien que les thèmes étudiés soient principalement géologiques, et non directement géographiques, ils sont également traités à l'aide de SIG. De nombreuses cartes sont produites grâce à des SIG-logiciels dans le but d'aider les planificateurs, aussi décideurs territoriaux, à prendre les décisions adéquates en matière de sauvegarde des ressources.

Les pays africains ne sont pas les seuls pays en voie de développement à avoir pris conscience des atouts de ces outils. Ces derniers sont, par exemple, également utilisés pour préparer des interventions d'urgence dans le bassin du *Mékong* (AUBE *et al.* 2006). Au Cambodge, en raison de la guerre civile ayant sévi jusqu'en 1991, les données à disposition étaient trop peu nombreuses et de qualité médiocre. Des nouvelles acquisitions de données s'avérèrent donc indispensables. Deux organismes canadiens souscrivirent un protocole de collaboration avec la *Mékong River Commission* (MRC) en vue de développer des instruments de géomatique et de télédétection utiles à l'amélioration de la gestion du fleuve ainsi qu'à une meilleure prise de décision en cas de débordement en zones habitées.

Le Vietnam, quant à lui, n'utilise pas un SIG mais bien un système d'information du territoire (SIT) dans la gestion de ses données cadastrales. Ce pays rencontre cependant de grosses difficultés à

développer son SIT étant donné que les règlements techniques, en vigueur là-bas, sont en retard sur celles des SIG en général et des SIT en particulier (DOI *et al.* 2009). Afin de pouvoir continuer à utiliser et tirer profit de son SIT, Hanoï a toutefois émis la volonté de normaliser, et ainsi rendre interopérables, les données.

Une application urbaine d'un SIG a été développée pour Antananarivo, capitale de Madagascar, afin de soutenir la lutte contre les risques naturels (MANDIMBIHARISON et RAHARISON, 2003). Cette ville est implantée dans une région de collines, dont certaines sont à forte pente et donc sensibles à l'érosion. Des glissements de terrain peuvent menacer les habitations. L'urbanisation incontrôlée des flancs de colline aggrave les phénomènes de ravinement et provoque le colmatage rapide des caniveaux ainsi que l'ensablement des fonds bas. D'autres zones très urbanisées se trouvent en fond de vallée et peuvent, par périodes de fortes pluies, être menacées de crues brutales et destructrices. Le SIG est un précieux outil d'aide à la gestion environnementale de la ville ainsi qu’aux prises de décision judicieuses. Des cartes d’érosion et de risques d'inondation ainsi qu’une carte générale de risques ont ainsi pu être dressées. Le SIG permet de gérer au mieux les ressources naturelles ainsi que l’urbanisation de la ville via l'élaboration d'un plan d'aménagement urbanistique.

Les exemples ci-dessus démontrent, à suffisance, l'importance de tels outils dans la gestion territoriale des pays en voie de développement. Si l'implantation de tels outils est de plus en plus réalisée dans les pays industrialisés, par contre, elle reste fortement limitée dans les pays en voie de développement principalement pour des raisons budgétaires. Ces outils devraient pourtant être utilisés bien davantage afin, entre autres, de réduire les coûts d'intervention en cas de catastrophe naturelle (AUBE *et al.* 2006).

1.3 Présentation du travail

Le **chapitre 1** introduit le sujet de notre travail. Nous y rappelons l’importance des catastrophes naturelles en zone urbaine, particulièrement dans les pays en voie de développement, et mettons en évidence l’intérêt et les avantages que présentent le développement et l’utilisation d’un système d’information géographique dans le cadre de la gestion de ces risques.

Les objectifs de la recherche menée dans le cadre d’une collaboration avec un étudiant de la Vrije Universiteit Brussel, Emmanuel Mba NYAMBOD, sont développés au **chapitre 2**. Nous y procédons à une mise en contexte du projet en explicitant les raisons du choix de la zone géographique qui servira de base au travail, en précisant son cadre, et en identifiant les catastrophes naturelles survenues dans un passé récent.

Nous analysons la démarche à suivre ainsi que les raisons pour lesquelles nous avons opté pour le développement d’un prototype SIG seulement.

Nous détaillons ensuite les étapes de son développement : analyse des données, préparation et normalisation des données suivant la norme ISO, conception du modèle conceptuel, passage au modèle logique et implémentation physique de données.

La contextualisation du projet fait l’objet du **chapitre 3**. Nous y abordons les démarches entreprises par NYAMBOD pour acquérir les données brutes que nous avons ensuite analysées. Nous y décrivons aussi la manière dont nous nous y sommes pris pour, dans la mesure du possible, valider les données de base et les structurer en paquetages. La présentation de l’hypothèse de travail et de la contrainte clôture ce chapitre.

Les 3 chapitres suivants concernent les différentes étapes de la conception du prototype de SIG. Elles sont schématisées à la figure 1

Le **chapitre 4** porte sur les différentes opérations d'analyse et de préparation des données : analyse de qualité et sélection des données afin de ne retenir que celles qui sont nécessaires pour atteindre les buts fixés lors de la contextualisation, constitution des classes de données qui sont traitées dans un but d'optimisation géométrique, attributaire et de référencement, renommage de ces classes ainsi que création du catalogue de données opérationnelles.

Les différentes étapes de la modélisation conceptuelle et logique des données sont traitées au **chapitre 5**. Nous y abordons successivement l'élaboration du modèle conceptuel de données, la justification du choix des associations et des cardinalités, la vérification du modèle, la traduction vers le modèle logique par la mise à forme des classes via ajout des attributs de jointure, et la construction des classes d'association. Une dernière complétion au niveau attributaire, pour la classe *Building*, est réalisée à ce stade étant donné que le modèle logique est nécessaire à la réalisation concrète des liens entre données intervenant dans ce traitement.

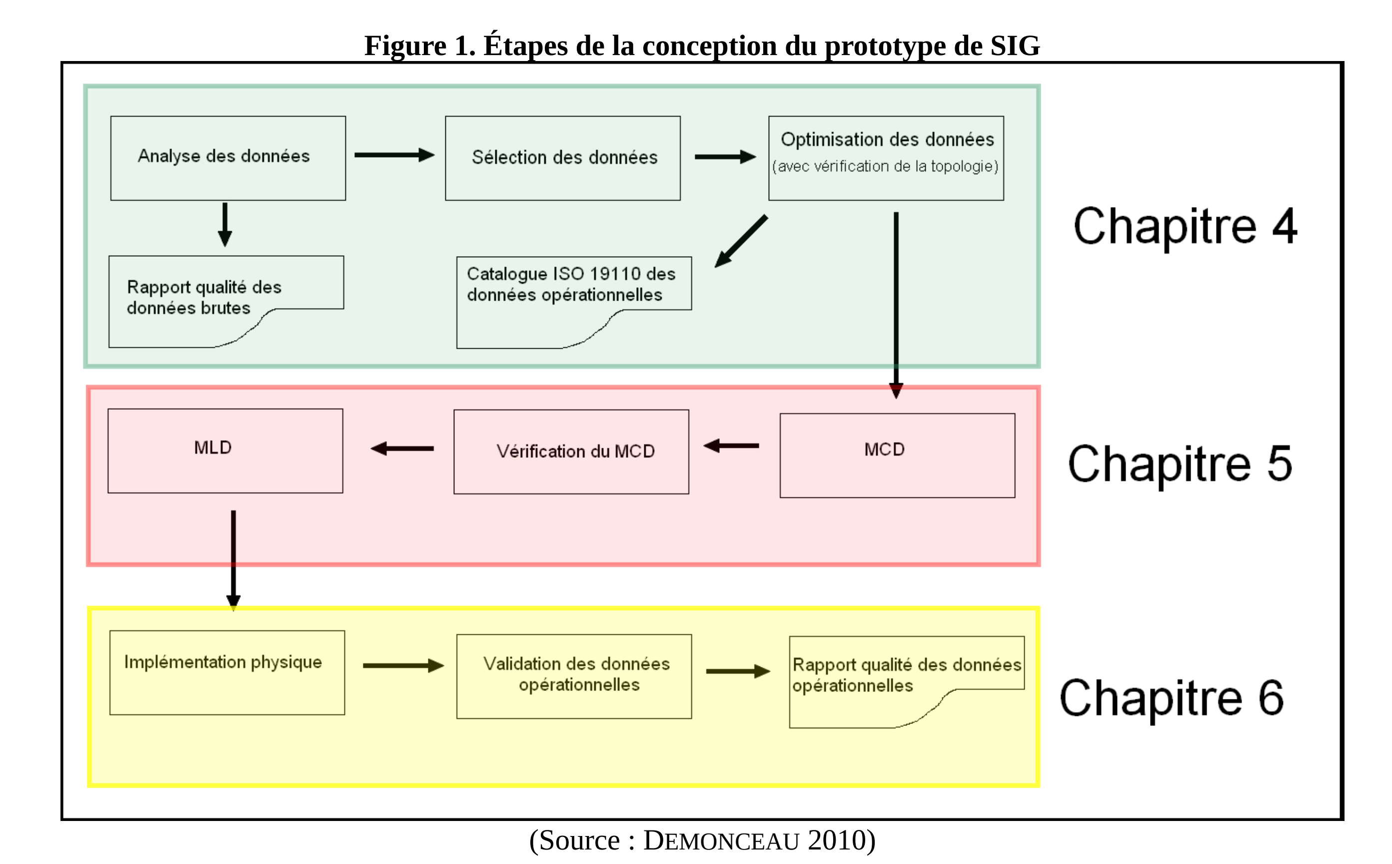
Le **chapitre 6** décrit l'implémentation physique du prototype après vérification du matériel informatique mis à disposition pour héberger la base de données. Cette base en question est créée et peuplée avec les données modélisées. Elle est rendue opérationnelle après validation des données opérationnelles via requêtes attributaires et spatiales ayant également la vocation de tester la consistance de la base. Le chapitre se clôt sur les manières possibles d'accéder à la base.

Le **chapitre 7** aborde le thème de l'exploitation du prototype. Cette exploitation, sortant du cadre du présent mémoire (en fait, elle fait partie de celui de NYAMBOD), se limitera ici à un exemple de visualisation des résultats d'une exploitation de notre prototype, quelques exemples d'enrichissement de la base, et un exemple d'exportation des données.

Notre conclusion est développée au **chapitre 8**. Les points originaux ainsi que les points faibles de l'entièreté du mémoire seront épinglées et décrites.

Le **chapitre 9** reprend les différentes annexes : le dictionnaire des classes brutes, la formule de conversion entre les coordonnées rectangulaires du système de projection UTM Zone 32 N et les coordonnées géodésiques WGS84 ainsi que, dans un volume séparé, le catalogue de données opérationnelles.

Figure 1. Étapes de la conception du prototype de SIG



(Source : DEMONCEAU 2010)

2 Objectifs de la recherche

2.1 Un travail en collaboration

2.1.1 Présentation générale

EMMANUEL NYAMBOD est étudiant en 2ième année du Master d'Écologie humaine à la *Vrije Universiteit Brussel* (VUB). Originaire de la ville de Bamenda au Cameroun, il a choisi de faire ses études en Belgique en raison de la qualité de l'enseignement dispensé. Le choix de son master découle de son souhait de mettre un jour ses connaissances et acquis au service de la collectivité de son pays d'origine et d'ainsi contribuer à améliorer la qualité de vie de ses concitoyens.

L'étude du lien corrélationnel entre l'urbanisation massive et incontrôlée de sa ville natale et l'augmentation de l'occurrence des risques naturels l'intéresse particulièrement. Il voudrait, via son mémoire, convaincre les autorités locales de la pertinence d'employer un SIG pour gérer les risques, résoudre certains problèmes urbains, et permettre aux décideurs territoriaux de prendre les justes décisions qui diminueront le nombre de pertes matérielles et humaines en cas de désastre. Il désire pouvoir mettre l'outil en ligne pour sensibiliser la population aux problèmes de risque, et lui permettre de communiquer ses idées et ses remarques en ce qui concerne le développement de la ville. Le choix de Bamenda, comme zone d'étude, est aussi lié au risque élevé d'occurrence de catastrophes naturelles telles que *flash floods*[1], inondations, glissements de terrain, coulées de boue, … (NYAMBOD 2010).

Son travail aurait été incomplet si l'aspect technique de la conception du SIG n'y avait été intégré. Or, il ne disposait personnellement d'aucune formation en matière technique et de traitement de données géographiques. Il a alors décidé d'externaliser ce volet. Dans cette démarche, il a contacté l'Université de Liège (ULg), seule institution universitaire en Belgique francophone à offrir une formation en Géomatique, dans le but d'y trouver une personne s'étant spécialisée dans cette matière, capable et d'accord de l'aider. Désirant relever le défi, car l'idée paraissait attrayante, nous nous sommes portés volontaires et avons répondu favorablement à cette proposition de collaboration.

Le présent mémoire fait donc suite au projet de collaboration entre :

- le demandeur, EMMANUEL NYAMBOD, qui possède la connaissance de terrain
- nous, futurs concepteurs du SIG, qui disposons de connaissances techniques à la réalisation d'un système facilitant les analyses.

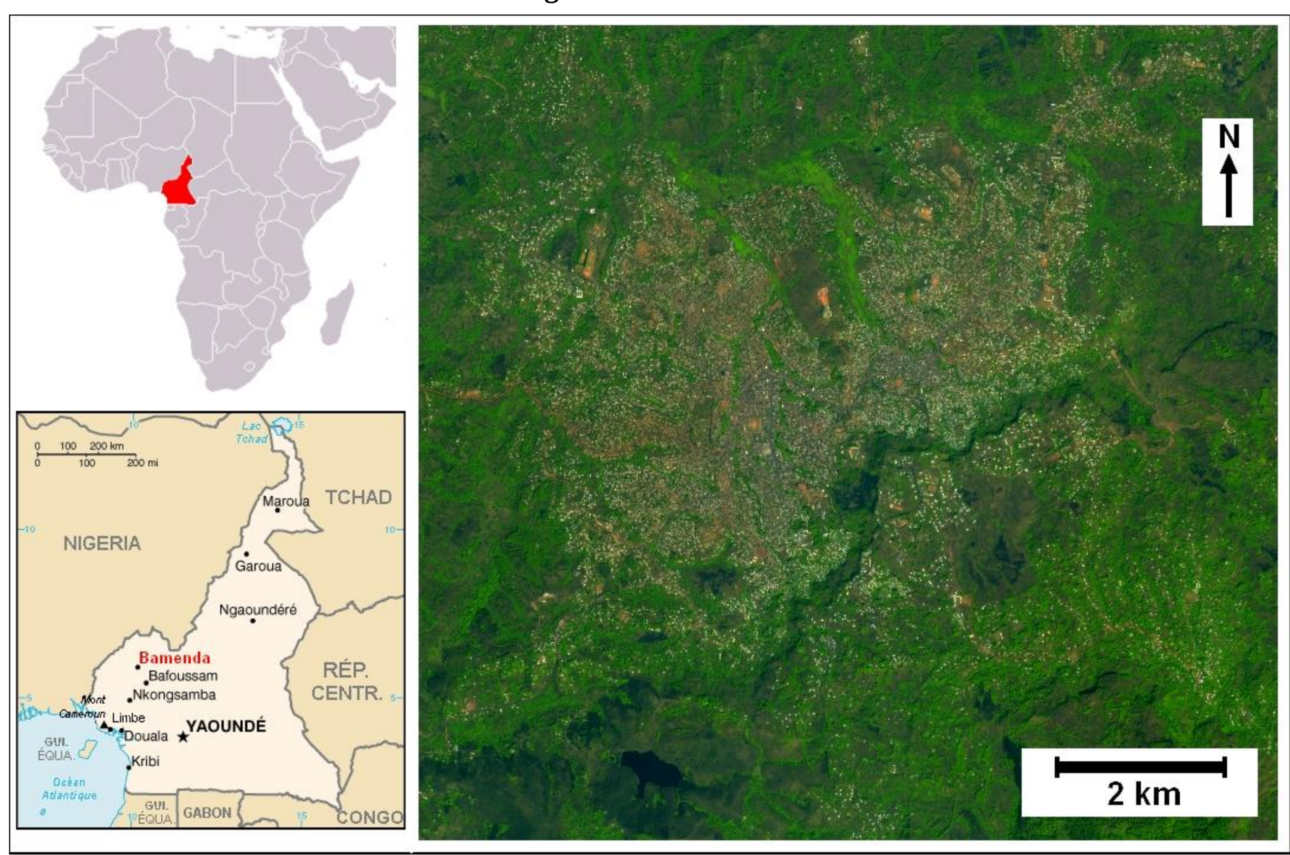
1 Crue subite (Source : http://www.reverso.net/text_translation.asp?lang=fr)

2.1.2 Mise en contexte

Bamenda est une ville du Cameroun localisée à environ 366 km au nord-ouest de la capitale Yaoundé (WIKIPEDIA 2010). Son centre-ville se situe aux coordonnées 10°08'51 E - 5°57'31 N. En 2007, sa population était estimée à 302.749 habitants (NYAMBOD 2010).

Le découpage administratif du Cameroun étant constitué de 10 régions qui regroupent plusieurs départements eux-mêmes composés d'arrondissements[2] (WIKIPEDIA 2010), Bamenda est une « *commune urbaine à régime spécial* » à la fois chef-lieu de la Région du Nord-Ouest, première ville du Département de Mezam, et à cheval sur trois arrondissements que sont : Bamenda Ier, Bamenda IIe et Bamenda IIIe.

Figure 2. Site d'étude



(Sources : sup. g. . : fr.wikipedia.org/wiki ; inf. d. : http://aceproject.org/regions-en/countries-and-territories/CM/copy_of_Carte_cameroun_francais.gif ; d. : Ikonos 2006)

Les catastrophes naturelles relativement fréquentes dans la ville sont de types divers (NYAMBOD 2010) : glissements de terrain, inondations, flash floods, etc.

Ces trois types de catastrophes sont plus fréquentes entre juillet et septembre, lors de la saison des pluies. La pluviométrie durant ces trois mois varie entre 340 et 450mm d'eau.

2 *Subdivisions* en anglais

Les glissements de terrain se produisent principalement le long de l'escarpement d'orientation sud-ouest/nord-est, et scindent la ville en deux grandes parties (figure 2). La partie supérieure, appelée *Up-station*, contient le siège d'administration de la ville tandis que la partie inférieure abrite la plus grande part de population. Le dénivelé entre partie haute et partie basse est en moyenne de 130m mais peut atteindre 200m par endroits (NYAMBOD 2010). La pente moyenne est d'environ 53 % (28 degrés). De nombreuses sources existent dans la partie supérieure de la ville et les cours d'eau qui se forment tant en surface qu'en profondeur coulent le long de cette pente, affaiblissent la cohésion des terres par leur apport d'humidité et établissent une zone d'instabilité. Au bas de la ville, l'urbanisation s'accompagne souvent d'excavations de terres préalables à la construction de fondations. Cela perturbe l'agencement des terres et entraîne des pertes de résistance qui fragilisent les parois de l'escarpement.

Tous ces facteurs, ainsi que la fragilité des parois due au sol essentiellement volcanique (MAKON MA PONDI 2009; TONYE & AKONO 2002; SIGAFRIQUE 2010), sont propices aux glissements de terrain. Le dernier glissement important en date a eu lieu le 4 août 2009, emportant la route principale reliant Bamenda à Baffoussam (MAKON MA PONDI 2009)(figure 3). FONYE précise même que l'escarpement est une zone écologique propice aux glissements de terrain principalement dus à des causes naturelles (vent, ensoleillement, pluies, ..) (dans NDULA & NDEFRU 2009) mais aussi humaines (pression et urbanisation anarchiques). NJOG, déclare dans un récit de voyage effectué dans la région, que ce genre de problème n'est pas unique, et rapporte les dires du préfet du département Mezam disant qu' « *il faut toujours user de tact pour calmer les populations à chaque fois qu'un glissement de terrain se produit* » (NJOG 2009).

MAKON MA PONDI s'interroge sur la qualité, souvent déficiente, de la gestion et de la maintenance des ouvrages d'art devant servir à limiter les risques (MAKON MA PONDI 2009). Il reproche un important degré de laxisme aux ingénieurs de l'association camerounaise des ingénieurs de maintenance (ACIM) qui prennent les décisions de renforcement des parois sujettes aux glissements de terrain et qui sont responsables des travaux de consolidation des structures vulnérables (MAKON MA PONDI 2009). L'auteur de cet article s'interroge fortement sur la destination inconnue des fonds budgétés pour la réalisation de ces travaux. NSEUMI déclare qu'« *il ne se passe plus un jour sans que l'on annonce des éboulements de terrains sur une partie du territoire* » (en parlant du territoire camerounais en général) (NSEUMI LEA 2009).

Figure 3. Inondation et glissement de terrain du 4 août 2009



(Source : NYAMBOD 2009)

La situation concernant les inondations n'est guère meilleure. Fréquemment, les débordements de rivières provoquent des dégâts aux habitations (ACHO-CHI 1998). Le même auteur signale que 20% des habitations en général se trouvent en zones inondables et que l'urbanisation incontrôlée se traduit par la construction de bâtiments en zones menacées. Les inondations menacent également les populations locales comme en témoigne l'exemple d'un indigène qui a déclaré avoir été victime d'un *flash flood* alors qu'il pêchait sur l'une des rivières de la ville (NYAMBOD 2010). Le risque d'inondation s'est accru suite à l'imperméabilisation des sols (routes, parkings...), la déforestation des collines, et l'écoulement des eaux perturbé par une forte charge en déchets de tous genres.

Les problèmes liés aux risques naturels dans la ville ont conduit NYAMBOD à s'intéresser de plus près à cette thématique, et à réaliser, via son mémoire, un certain nombre de simulations dans le but de cartographier des solutions à proposer aux décideurs locaux.

2.1.3 Conception du SIG

Le choix de la démarche dans la conception d'un SIG dépend à la fois de l'incertitude du projet et de sa complexité (voir figure 4).

Figure 4. Choix de la démarche

Complexité du projet		Incertitude du projet : Faible	Incertitude du projet : Élevée
	Élevée	Démarche traditionnelle (ex. MÉRISE, cascade)	Démarche traditionnelle avec prototypage (ex. mixte, évolutive)
	Faible	Prototypage seulement (utilisation sans contexte méthodologique)	Prototypage seulement (utilisation sans contexte méthodologique)

(Source : DONNAY 2009)

La complexité s'évalue sur base des critères suivants (DONNAY 2009) :
- le nombre d'utilisateurs ;
- la quantité d'informations nouvelles introduites ;
- la complexité des informations elles-mêmes.

Dans notre cas, il y a un seul utilisateur et les informations sont peu complexes car elles sont déjà sous un format utilisable par les logiciels de gestion de base de données. Par contre, la quantité d'informations aurait été très élevée si l'on avait pris en compte le grand nombre de risques énumérés par NYAMBOD. Mais pour rester dans les limites d'un travail de fin d'études, il a été décidé de réduire à deux le nombre de risques et, de ce fait, la quantité d'informations. Nous n'avons retenu que les glissements de terrain et les inondations qui, d'après l'étudiant camerounais, constituent les risques naturels les plus préoccupants dans le cas de la ville de Bamenda.

L'incertitude s'évalue sur base des critères suivants (DONNAY 2009) :

- le niveau de structuration de l'outil déjà implémenté ;
- la prise de conscience par l'utilisateur de son rôle à jouer dans les étapes de la conception et la compréhension par le concepteur des besoins de l'utilisateur ;
- l'expérience du concepteur.

Le premier critère ne s'applique pas étant donné qu'il n'existe actuellement aucun outil sur lequel nous pourrions nous appuyer. Le niveau d'incertitude, selon le second critère, est élevé étant donné que l'utilisateur NYAMBOD n'a aucune formation dans le domaine et que sa formulation des besoins est imprécise et variable. L'expérience du concepteur, en troisième critère, est ici telle qu'il ne dispose que de connaissances théoriques et peu de connaissances pratiques.

En résumé, dans notre cas, la complexité du projet étant faible et son incertitude élevée, nous n'avons pas opté pour la démarche traditionnelle mais pour le prototypage seulement. Un prototype est par définition, un *« logiciel intégrant tout ou partie des fonctionnalités réelles à développer mais n'utilisant qu'un sous-ensemble de données réelles en entrée ; la technique de développement utilisée sera celle retenue pour l'ensemble du système, ce point de vue étant celui du réalisateur »* (DONNAY 2009)

Dans la conception de ce prototype, nous n'avons conservé que les données relatives à la vulnérabilité (des infrastructures, par exemple) et aux aléas (AYRAL *et al.* 2001).

Le prototype est conçu sur base d'une approche systémique où « *l'organisation est au centre de la méthode et non le projet* » (DONNAY 2009). Tout au long de la conception du prototype et à chaque étape de sa réalisation, une collaboration étroite et une communication de qualité ont été indispensables entre les « *concepteurs* » que nous sommes et le « *client* » étudiant à la VUB, afin de rencontrer les demandes précises que ce dernier formule (LAURINI 2001). Il incombait donc à NYAMBOD d'établir, dès le début du projet, son cahier des charges, c'est-à-dire la description aussi complète que possible des fonctionnalités que doit remplir le prototype.

L'état de l'art a mis en évidence le fait que les administrations de la plupart des pays en voie de développement n'ont pas les moyens d'adopter l'outil SIG. C'est également le cas de la ville de Bamenda. Le matériel informatique en place ne se prête pas toujours au déploiement d'un SIG. Certains bureaux ne disposent d'aucun ordinateur. D'autres n'en possèdent qu'un, et encore il s'agit d'un ancien modèle. Lorsqu'enfin un ordinateur en état de marche est disponible, il y a souvent un manque de formation du personnel à l'utilisation des logiciels SIG de bureau (NYAMBOD 2010). Le budget de la ville consacré au développement informatique, et plus spécifiquement d'un SIG, est quasiment inexistant, l'équipement étant très cher et la ville ayant d'autres priorités d'investissement.

Au vu de ces constatations et dans le cadre de ce mémoire, il est évident que l'option de mise en place d'un système important et onéreux ainsi que d'utilisation de logiciels payants ne peut être prise. De la conception à l'exploitation, les logiciels utilisés seront tous de licence *Open Source*, logiciels gratuits développés par une communauté de concepteurs.

2.2 Conception du SIG sur le territoire considéré selon standards et couts minimaux

2.2.1 Transformation des données existantes

Les données acquises pour réaliser ce travail feront, en premier lieu, l'objet d'une analyse approfondie pour évaluer leur qualité globale. LAURINI distingue (dans LAURINI 2001) deux types de contrôle de qualité :

- le contrôle-qualité mettant en œuvre des techniques et actions susceptibles de vérifier que les standards de qualité sont rencontrés ;
- l'estimation-qualité permettant d'évaluer le degré de qualité des programmes et procédés pratiques mis en place, ainsi que celui des produits et services liés au prototype.

L'auteur énumère plusieurs aspects à prendre en considération lors de l'évaluation de la qualité des données :

- l'origine, la précision, la résolution, la complétude et le degré de consistance des données ;
- la qualité des métadonnées ;
- le coût de la qualité et de la maintenance.

Les données brutes feront donc l'objet d'une analyse dans chacun des deux domaines cités par LAURINI, et cela pour chaque aspect qualité mentionné par lui. La fiabilité des données brutes sera estimée sur base du rapport de qualité établi en fin d'analyse.

NYAMBOD nous a cependant déjà avertis que l'entreposage dans des tiroirs et armoires des données stockées sur disquettes et CD-ROM avait été très précaire et qu'il permettait de présager un manque de structures entre elles.

Une sélection des donnés utiles, réalisée en commun avec NYAMBOD afin d'intégrer les points de vue du concepteur et de l'utilisateur, a permis de ne garder que les données ayant un lien direct avec les deux thématiques abordées. Les données utiles retenues peuvent être classifiées selon trois facteurs : risque, physique et humain.

Le facteur « risque » englobe les données relatives aux facteurs déclencheurs d'inondations et/ou de glissements de terrain.

Le facteur « physique » concerne les pertes financières engendrées par l'endommagement des structures construites.

Le facteur « humain » est relatif, quant à lui, à la population, et plus spécialement à la localisation des habitations. Ce dernier facteur est d'une importance capitale pour estimer les pertes humaines en cas de catastrophe dans l'applicatif de NYAMBOD.

Les données sélectionnées ont été préparées en corrigeant, autant que faire se peut, les défauts identifiés dans le rapport de qualité des données brutes. Cette préparation a eu pour but d'optimiser les données en ne gardant que les informations les plus utiles, et en supprimant les redondances. La normalisation des données attributaires a été réalisée sur base de la norme ISO 19110, norme reconnue qui traite du catalogage des données. Le but était de rendre les données interopérables et faciliter les échanges entre services (ici, en l'occurrence, entre concepteur et client) en fournissant une description

claire, normalisée et non ambiguë du contenu du prototype. Cette norme a notamment été utilisée pour réaliser le catalogue des données du Cadastre du Vietnam (DOI *et al* 2009), ou celui de l'Institut Géographique National (IGN) dans son dictionnaire *Table Of Contents* (TOC), ou encore celui de la Compagnie intercommunale liégeoise des eaux (CILE) (COLLIGNON 2004).

Les outils utilisés pour l'analyse et la transformation des données sont de licence *Open Source* afin de rencontrer la contrainte de coût minimal. Nous avons uniquement utilisé *OpenJump*, SIG-logiciel codé en Java, pour les opérations de nettoyage de données, de corrections de géométries et d'attributs.

2.2.2 Modèle conceptuel de données

Il existe en général différents niveaux dans la modélisation d'un système d'informations :

- le premier est le niveau externe (LAURINI 2001). Il s'agit de la toute première représentation du monde extérieur et constitue déjà une étape dans la conception de la base de données qu'est le prototype dans notre cas. Le niveau externe est celui de la découverte des données et des interprétations que chaque observateur en fait. Dans notre cas, les modèles externes sont au nombre de deux : la vue qu'a le concepteur et la vue qu'a le demandeur (LAURINI 2001). Il ne faut pas, en effet, oublier que la sélection des données a été réalisée en commun accord et que chacun y a apporté ses visions (CALKINS 1997) ;
- le deuxième est le niveau conceptuel qui synthétise tous les modèles externes. Le modèle conceptuel permet d'offrir une description du contenu de la future base de données ainsi que des relations qui existent entre les données ;
- le troisième est le niveau logique ;
- le quatrième est le niveau physique.

Ces deux derniers niveaux seront abordés dans la sous-section 2.2.3.

Pour l'élaboration du schéma conceptuel, nous avons opté pour le langage de modélisation *Unified Modeling Language* (UML). Il s'agit d'un langage standardisé de description des états statique et dynamique d'un système orienté « objet » (BOOCH *et al.* 1999). Dans notre cas, le système est un prototype.
Les modèles UML sont exploités pour :

- la compréhension des données et de leurs traitements ;
- la réflexion quant à la manière de réaliser le futur produit informatique qui implémente ces modèles ;
- une communication non ambiguë, entre informaticiens et utilisateurs, d'informations sur le système (DONNAY 2005).

Le modèle conceptuel est, en théorie, indépendant des outils informatiques qui sont utilisés pour l'implémenter. Il existe neuf diagrammes au sein d'UML mais seul le diagramme de classes sera pris en considération ici, étant donné que c'est uniquement la structure statique de la base de données qui nous intéresse.

Un pictogramme évocateur de la primitive géométrique des données sera ajouté à la représentation des classes (LAURINI 2001) pour tenir compte de la spatialité des données car le formalisme UML n'est conçu que pour fournir une description générale (et non spécifique) des objets géographiques
Le formalisme UML a été utilisé plusieurs fois avec succès dans des travaux antérieurs de réalisation de schémas conceptuels, notamment par LAPLANCHE qui fut convaincu de ses performances (dans

LAPLANCHE 2001 & 2002), et par BALRAM et DRAGICEVIC, qui modélisèrent des schémas de collaboration (BALRAM & DRAGICEVIC 2006).

Le schéma conceptuel sera créé sous *ArgoUML*, petit logiciel codé en Java et disponible dans sa version la plus récente qui est la version 0.30.1. Nous ne disposons cependant pas d'un atelier de génie logiciel[3] (AGL) supporté par un logiciel *Open Access* qui nous permette de passer du modèle conceptuel au modèle logique. Cette phase devra être réalisée manuellement (avec *OpenJump* ici).

2.2.3 Implémentation d'un prototype en *OpenAccess*

D'une manière pratique, l'implémentation consiste à passer du modèle conceptuel au modèle physique mais cela ne peut pas être réalisé sans l'étape intermédiaire de description du modèle logique. Cette dernière consiste à identifier la manière dont seront réalisés les liens qui existeront au niveau physique. La réalisation de liens entre les données est possible grâce aux « attributs de jointures ». Ceux-ci ne sont pas nécessairement présents dans les tableaux associés aux classes à implémenter, et lorsqu’ils en sont absents, il faut les ajouter en effectuant des traitements sur les données.

La base de données « test » à partir de laquelle seront effectués les essais de conception est créée grâce au Système de Gestion de Bases de Données (SGBD) *PostgreSQL* complété avantageusement par la cartouche *PostGIS* qui permet de tenir compte de la spatialité des données manipulées. Une interface graphique *PgAdmin III* a également été installée pour faciliter la manipulation de la base *PostGIS*. Elle est d'une aide incontestable pour la création de la base de données, le chargement des données et l'exploitation finale de l'outil. Sans elle, le temps de conception serait considérablement allongé étant donné que toutes les requêtes devraient être encodées manuellement en lignes de commande.

Ces trois modules ont été installés en une seule étape en téléchargeant le *StackBuilder* disponible sur le site officiel de *PostgreSQL* (http://www.postgresql.org/). Après avoir démontré l'efficacité du poste hôte, le prototype a été implémenté sur le serveur B5 de l'Université de Liège. Ce serveur héberge déjà quelques bases de données (il ne devait donc pas y avoir de problème à en ajouter une nouvelle) ainsi qu’un SGBD *PostGIS* tournant sous l'*Operating System* (OS) *Kubuntu* qui est une distribution GNU[4]/Linux.

La gratuité de cet OS et de ses mises à jour, ainsi que la gratuité des logiciels *ArgoUML*, pour la modélisation conceptuel, et *OpenJump* utilisé principalement comme *viewer* au même titre que *Quantum GIS (Qgis)* et *User-friendly Desktop Internet GIS (uDig)*, nous permettent de satisfaire à la contrainte de minimisation des coûts. Il existe, bien sûr, bon nombre d'autres logiciels de la même gamme qui auraient pu être utilisés (*Géomatica*, *AvisMap Free*, ...), mais nous avons opté pour ceux que nous jugions les plus complets et les plus fiables.

3 Outil d'aide à la conception portant surtout sur la transformation des schéma, la production de l'entièreté de la documentation du SI et la validation de solutions possibles par la conception de prototypes

4 GNU's Not Unix

2.3 Conception du Projet

Le projet de collaboration a un impact tant sur le mémoire de NYAMBOD que sur le nôtre. Il s'agit de deux mémoires développés en parallèle, et pour lesquels les échanges entre acteurs ont dû être nombreux afin de mener à bien cette collaboration. L'assignation du temps nécessaire à la réalisation de chacune des étapes du développement de notre prototype a donc dû tenir compte de la vitesse d'avancement de chacun dans la réalisation de ses tâches respectives.

De notre côté, nous étions surtout disponibles de février à mai, le premier quadrimestre ayant été principalement consacré aux cours et travaux pratiques. Nous avons estimé nécessaire d'adopter, dès février, un rythme de travail de 8 heures par jour et 5 jours par semaine (week-end compris) afin d'avancer suffisamment rapidement dans la réalisation de notre travail (la *dead line* ayant été fixée à mi-mai pour parer aux éventuels imprévus) tout en assistant aux quelques cours et travaux pratiques du second quadrimestre.

Le sujet de mémoire nous a été attribué fin juin 2009 et les mois de juillet, août et septembre ont été consacrés à la lecture de la littérature portant sur le sujet.

L'analyse des données acquises a été réalisée durant l'entièreté du premier quadrimestre (de septembre à décembre) et finalisée, après les examens, en fin janvier, avec la réalisation d'un dictionnaire de données brutes (annexe 1).

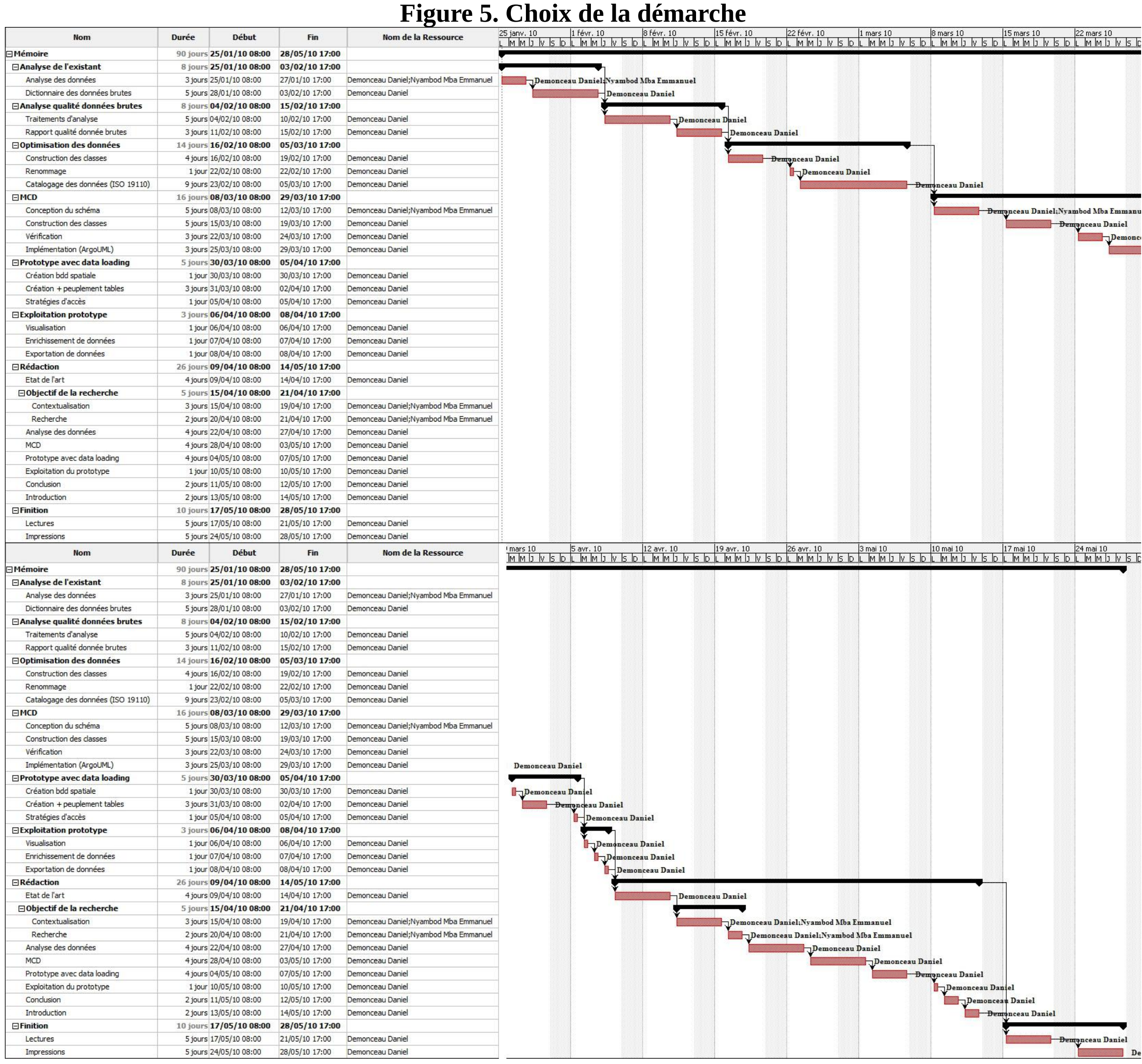
Dans ce calendrier, un strict minimum d'un mois a été attribué à la rédaction de l'ensemble du mémoire et une semaine complète à sa relecture et à son impression.

De son côté, NYAMBOD a débuté son travail au mois de juillet. Durant ce mois-là, il a acheté les données et contacté les responsables des trois Administrations de la ville de Bamenda ainsi que les délégations provinciales responsables de l'aménagement du territoire de la Région du Nord-Ouest au Cameroun. A la fin de ce même mois, il a mené une enquête en distribuant des questionnaires et en menant des interviews, et a mis à jour les données achetées. Il est ensuite revenu en Europe pour interpréter les résultats issus de ces opérations et pour nous permettre de progresser dans l'avancement de la conception du prototype.

L'estimation du temps nécessaire à la réalisation des différentes étapes du travail est illustrée dans le Gantt[5] repris à la figure 5:

5 Outil utilisé en ordonnancement et gestion de projet (WIKIPEDIA 2010)

Figure 5. Choix de la démarche



Nom	Durée	Début	Fin	Nom de la Ressource
⊟Mémoire	90 jours	**25/01/10 08:00**	**28/05/10 17:00**	
⊟Analyse de l'existant	8 jours	**25/01/10 08:00**	**03/02/10 17:00**	
Analyse des données	3 jours	25/01/10 08:00	27/01/10 17:00	Demonceau Daniel;Nyambod Mba Emmanuel
Dictionnaire des données brutes	5 jours	28/01/10 08:00	03/02/10 17:00	Demonceau Daniel
⊟Analyse qualité données brutes	8 jours	**04/02/10 08:00**	**15/02/10 17:00**	
Traitements d'analyse	5 jours	04/02/10 08:00	10/02/10 17:00	Demonceau Daniel
Rapport qualité donnée brutes	3 jours	11/02/10 08:00	15/02/10 17:00	Demonceau Daniel
⊟Optimisation des données	14 jours	**16/02/10 08:00**	**05/03/10 17:00**	
Construction des classes	4 jours	16/02/10 08:00	19/02/10 17:00	Demonceau Daniel
Renommage	1 jour	22/02/10 08:00	22/02/10 17:00	Demonceau Daniel
Catalogage des données (ISO 19110)	9 jours	23/02/10 08:00	05/03/10 17:00	Demonceau Daniel
⊟MCD	16 jours	**08/03/10 08:00**	**29/03/10 17:00**	
Conception du schéma	5 jours	08/03/10 08:00	12/03/10 17:00	Demonceau Daniel;Nyambod Mba Emmanuel
Construction des classes	5 jours	15/03/10 08:00	19/03/10 17:00	Demonceau Daniel
Vérification	3 jours	22/03/10 08:00	24/03/10 17:00	Demonceau Daniel
Implémentation (ArgoUML)	3 jours	25/03/10 08:00	29/03/10 17:00	Demonceau Daniel
⊟Prototype avec data loading	5 jours	**30/03/10 08:00**	**05/04/10 17:00**	
Création bdd spatiale	1 jour	30/03/10 08:00	30/03/10 17:00	Demonceau Daniel
Création + peuplement tables	3 jours	31/03/10 08:00	02/04/10 17:00	Demonceau Daniel
Stratégies d'accès	1 jour	05/04/10 08:00	05/04/10 17:00	Demonceau Daniel
⊟Exploitation prototype	3 jours	**06/04/10 08:00**	**08/04/10 17:00**	
Visualisation	1 jour	06/04/10 08:00	06/04/10 17:00	Demonceau Daniel
Enrichissement de données	1 jour	07/04/10 08:00	07/04/10 17:00	Demonceau Daniel
Exportation de données	1 jour	08/04/10 08:00	08/04/10 17:00	Demonceau Daniel
⊟Rédaction	26 jours	**09/04/10 08:00**	**14/05/10 17:00**	
Etat de l'art	4 jours	09/04/10 08:00	14/04/10 17:00	Demonceau Daniel
⊟Objectif de la recherche	5 jours	**15/04/10 08:00**	**21/04/10 17:00**	
Contextualisation	3 jours	15/04/10 08:00	19/04/10 17:00	Demonceau Daniel;Nyambod Mba Emmanuel
Recherche	2 jours	20/04/10 08:00	21/04/10 17:00	Demonceau Daniel;Nyambod Mba Emmanuel
Analyse des données	4 jours	22/04/10 08:00	27/04/10 17:00	Demonceau Daniel
MCD	4 jours	28/04/10 08:00	03/05/10 17:00	Demonceau Daniel
Prototype avec data loading	4 jours	04/05/10 08:00	07/05/10 17:00	Demonceau Daniel
Exploitation du prototype	1 jour	10/05/10 08:00	10/05/10 17:00	Demonceau Daniel
Conclusion	2 jours	11/05/10 08:00	12/05/10 17:00	Demonceau Daniel
Introduction	2 jours	13/05/10 08:00	14/05/10 17:00	Demonceau Daniel
⊟Finition	10 jours	**17/05/10 08:00**	**28/05/10 17:00**	
Lectures	5 jours	17/05/10 08:00	21/05/10 17:00	Demonceau Daniel
Impressions	5 jours	24/05/10 08:00	28/05/10 17:00	Demonceau Daniel

Nom	Durée	Début	Fin	Nom de la Ressource
⊟Mémoire	90 jours	**25/01/10 08:00**	**28/05/10 17:00**	
⊟Analyse de l'existant	8 jours	**25/01/10 08:00**	**03/02/10 17:00**	
Analyse des données	3 jours	25/01/10 08:00	27/01/10 17:00	Demonceau Daniel;Nyambod Mba Emmanuel
Dictionnaire des données brutes	5 jours	28/01/10 08:00	03/02/10 17:00	Demonceau Daniel
⊟Analyse qualité données brutes	8 jours	**04/02/10 08:00**	**15/02/10 17:00**	
Traitements d'analyse	5 jours	04/02/10 08:00	10/02/10 17:00	Demonceau Daniel
Rapport qualité donnée brutes	3 jours	11/02/10 08:00	15/02/10 17:00	Demonceau Daniel
⊟Optimisation des données	14 jours	**16/02/10 08:00**	**05/03/10 17:00**	
Construction des classes	4 jours	16/02/10 08:00	19/02/10 17:00	Demonceau Daniel
Renommage	1 jour	22/02/10 08:00	22/02/10 17:00	Demonceau Daniel
Catalogage des données (ISO 19110)	9 jours	23/02/10 08:00	05/03/10 17:00	Demonceau Daniel
⊟MCD	16 jours	**08/03/10 08:00**	**29/03/10 17:00**	
Conception du schéma	5 jours	08/03/10 08:00	12/03/10 17:00	Demonceau Daniel;Nyambod Mba Emmanuel
Construction des classes	5 jours	15/03/10 08:00	19/03/10 17:00	Demonceau Daniel
Vérification	3 jours	22/03/10 08:00	24/03/10 17:00	Demonceau Daniel
Implémentation (ArgoUML)	3 jours	25/03/10 08:00	29/03/10 17:00	Demonceau Daniel
⊟Prototype avec data loading	5 jours	**30/03/10 08:00**	**05/04/10 17:00**	
Création bdd spatiale	1 jour	30/03/10 08:00	30/03/10 17:00	Demonceau Daniel
Création + peuplement tables	3 jours	31/03/10 08:00	02/04/10 17:00	Demonceau Daniel
Stratégies d'accès	1 jour	05/04/10 08:00	05/04/10 17:00	Demonceau Daniel
⊟Exploitation prototype	3 jours	**06/04/10 08:00**	**08/04/10 17:00**	
Visualisation	1 jour	06/04/10 08:00	06/04/10 17:00	Demonceau Daniel
Enrichissement de données	1 jour	07/04/10 08:00	07/04/10 17:00	Demonceau Daniel
Exportation de données	1 jour	08/04/10 08:00	08/04/10 17:00	Demonceau Daniel
⊟Rédaction	26 jours	**09/04/10 08:00**	**14/05/10 17:00**	
Etat de l'art	4 jours	09/04/10 08:00	14/04/10 17:00	Demonceau Daniel
⊟Objectif de la recherche	5 jours	**15/04/10 08:00**	**21/04/10 17:00**	
Contextualisation	3 jours	15/04/10 08:00	19/04/10 17:00	Demonceau Daniel;Nyambod Mba Emmanuel
Recherche	2 jours	20/04/10 08:00	21/04/10 17:00	Demonceau Daniel;Nyambod Mba Emmanuel
Analyse des données	4 jours	22/04/10 08:00	27/04/10 17:00	Demonceau Daniel
MCD	4 jours	28/04/10 08:00	03/05/10 17:00	Demonceau Daniel
Prototype avec data loading	4 jours	04/05/10 08:00	07/05/10 17:00	Demonceau Daniel
Exploitation du prototype	1 jour	10/05/10 08:00	10/05/10 17:00	Demonceau Daniel
Conclusion	2 jours	11/05/10 08:00	12/05/10 17:00	Demonceau Daniel
Introduction	2 jours	13/05/10 08:00	14/05/10 17:00	Demonceau Daniel
⊟Finition	10 jours	**17/05/10 08:00**	**28/05/10 17:00**	
Lectures	5 jours	17/05/10 08:00	21/05/10 17:00	Demonceau Daniel
Impressions	5 jours	24/05/10 08:00	28/05/10 17:00	Demonceau Daniel

(Source : DEMONCEAU 2009)

3 Contextualisation

3.1 Démarche externe Emmanuel

Pour nous permettre de réaliser nos objectifs, et dans le cadre de la collaboration VUB-ULg, nous avons acquis les données numériques en juillet 2009 auprès de l'Administration générale de la ville de Bamenda. Ces données ont été produites par Cameroon Geographic Information Systems (CamGIS Ltd), petite entreprise camerounaise créée en 2001 et spécialisée dans la géo-information. La rentabilisation de cet investissement doit passer par l'exploitation cohérente et utile des données afin de rencontrer les demandes formulées par NYAMBOD dans son cahier des charges.

NYAMBOD a réalisé un inventaire des catastrophes naturelles récentes ayant affecté la ville de Bamenda ainsi que des dégâts qui y ont été causés.

Figure 6. Localisation, dégâts et date de certaines catastrophes naturelles

Landslide and mud flow		
Quarters	**Damages**	**Year**
Up-Station	Traffic arrest on major high way for about 24 hours	4th/08/2009
Sisia	One death, two seriously injured	4th/08/2009
Abangoh	Three deaths	2007
Subsidence		
Anindoh	Collapse of bridge on major national highway	7th/08/2009
Floods		
Mulang	Two death	2006
Ngomgham	Houses and properties	Annual
Nitop I	Houses and properties	Annual
Ntamulung, II and III	Houses and properties	Annual
Ntaturu	One death	2007
Mougheb	Swept away 700 bags of 50 kg rice, One death	2007
Azire	Houses and properties	Annual

(Source : NYAMBOD 2009)

Cet inventaire des catastrophes de passé récent confirme le choix des thématiques retenu comme base du développement du prototype de SIG : glissements de terrain et inondations.

En parallèle, il a adressé des questionnaires aux responsables de la ville afin de connaître l'état actuel des infrastructures SIG.

Il s'est également rendu sur le terrain pour recueillir, de vive voix, l'opinion et l'avis des habitants sur les risques naturels qu'encourt leur ville. Cela lui a permis de mieux cibler leurs besoins et d'orienter ses recherches en ce sens. Rappelons que le but ultime de son projet est l'élaboration d'une carte visualisant les zones les plus menacées par les risques.

Le prototype développé permettra de vérifier son hypothèse de départ à savoir qu'un SIG est utile à résoudre certains problèmes urbains identifiés grâce aux réponses aux questionnaires, aux interviews réalisées et au prototype développé.

3.2 **Présentation des données**

Les données numériques acquises consistent en un set de 70 *Shapefiles* (shp), une trentaine de fichiers *ArcView Theme Legend File* (avl), et une image aérienne orthorectifiée au format *Tagged Image File Format* (tiff).

Ces données sont constituées d'un ensemble de classes d'objets géographiques et sont non-structurées. Elles ont été fournies « en vrac » et comprennent de nombreuses informations redondantes. Elles varient tant par le niveau de leur précision et/ou exactitude géométriques que par les valeurs des attributs. Les primitives géométriques employées pour la représentation des objets géographiques sont de type ponctuel, linéaire, polygonal, poly-linéaire ou multi-polygonal.

Elles ont été digitalisées sur base des objets identifiables sur l'orthoimage. Certaines données relatives à différents domaines : réseau routier, hydrographie, relief, etc., ont été revues et mises à jour par NYAMBOD et le nom des fichiers retouchés portent le suffixe 0707. Pour une meilleure organisation ultérieure, et pour plus de clarté, une classification en paquetage a été réalisée sur base du caractère sémantique des données (LAPLANCHE 2006) (voir figure 7).

L'orthoimage, qui n'est intégrée à aucun paquetage, est le produit de l'orthorectification d'une image brute prise par le satellite *Ikonos* en février de l'année 2006 (jour et date précis inconnus). Elle couvre une zone sensiblement carrée de 10km de côté et centrée sur le centre-ville de Bamenda. Ses métadonnées sont analysées en sous-section 4.2.1.

Toutes les données du paquetage « *Administratif* » ont un lien plus ou moins étroit avec le découpage du territoire en zones administratives. Il y a parmi ces données des couches relatives aux quartiers (sous forme polygonale ou ponctuelle). Une couche contient également des informations sur les villages « *composant* »[6] la ville de Bamenda. Certaines couches reprennent, sous forme ponctuelle, les centres des villes voisines (*bamenda satellite towns*, *evolving hierarchy of settlements* et *subregional context_merge*). Les arrondissements de la ville sont reprises dans les classes débutant par « *subdiv* ». D'autres couches contiennent des polygones ne correspondant à aucun découpage réel du territoire (*evolution*, *quartier polygon*, *internal urban spatial structure* et *gradient analysis*). Une seule couche, enfin, contient des polygones de forme circulaire illustrant les centres d'affaires (*cbd*). Il a fallu être attentif à la répétition des couches « *semblables* » et identifier celles qui sont le plus représentatives de la réalité telles les couches digitalisées le plus récemment et contenant donc des informations actualisées.

---

6 Villes et villages se trouvent au même niveau hiérarchique dans le découpage administratif du Cameroun

Figure 7. Classification en paquetage des classes

Paquetage	*Classes*	Primitive géom.	Identifiant de classe
Administratif	bamenda satellite towns	Point	2
	cbd	Polygone	6
	evolution	Polygone	18
	evolving hierarchy of settlements	Point	19
	gradient analysis	Polygone	21
	internal urban spatial structure	Polygone	30
	limit urban area bamenda	Ligne	38
	limit urban area bamenda line	Ligne	39
	limite arrondissement	Polygone	40
	line covering sat towns	Ligne	41
	quarters mendankwe	Point	48
	quarters_bamenda	Point	49
	quarters_mendankwe	Point	50
	quarters0_0707	Point	51
	quartier polygon	Polygone	52
	subdivibdapolygon0_0707	Polygone	63
	subdivibdapolyline0_0707	Ligne	64
	subdivision line	Ligne	65
	subregional context_merge	Point	66
Affectation de sol	hydro bda lakes	Polygone	26
	landuse	Polygone	32
	landuse points	Polygone	33
	landuse points1207	Point	34
	landuse polygons 1207	Polygone	35
	landuse_study area	Polygone	36
	swamps	Polygone	67
Buildings	Catchment sec general and	Polygone	5
	council halls	Polygone	8
	eauip marchands	Point	12
	education	Point	13
	education0_0707	Point	14
	equipe public et parapublic	Point	15
	equipements d'hôtellerie et touristiques	Point	16
	equipements marchandes	Point	17
	finance	Point	20
	habitatcub0_0207	Polygone	22
	houses within catchment zone	Polygone	25
	health	Point	23
	hotel de ville		24
	industry	Point	29
	land marks	Point	31
	nursery school travelling govt	Polygone	42
	population figures	Point	45
	primary catchment zone	Polygone	46
	religious institution	Point	53
	securité publique	Point	60
	service public et parapublic	Point	61
	télécommunication	Point	68
	travelling distance sec school	Polygone	70
Hydrographie	hydro bda rivers	Ligne	27
	hydro_satellite towns	Ligne	28
	limit satillite town area	Ligne	37
	rivers sat towns	Ligne	54
Relief	contour satellite towns	Ligne	7
	courbes de niveau bamenda	Ligne	9
	courbes de niveau bamenda_urban area	Ligne	10
	station escarpment	Ligne	62
Voiries	artirary road buffer	Polygone	1
	bridges and culverts	Point	3
	buffer round critical junctions	Polygone	4
	critical areas and trouble spots	Point	11
	parking	Polygone	43
	parking facilities revised	Point	44
	primary road buffer	Polygone	47
	road accessibility	Polygone	55
	road network	Ligne	56
	roads linkin the evolving hieracchy of satellite towns	Ligne	57
	roads_0707	Ligne	58
	secondary road buffer	Polygone	59
	tertary road buffer	Polygone	69

(Source : DEMONCEAU 2010)

Le paquetage « *Affectation de sol* » reprend les classes de suffixe « *landuse* ». Nous y avons ajouté les classes relatives aux lacs et aux zones humides, respectivement *hydro bda lakes* et *swamps,* étant donné que ces dernières font partie des objets repris par les autres classes, et qu'ils ont donc un lien étroit. Les classes polygonales de *landuse* sont relativement complètes car elles contiennent des polygones représentant des zones de toute nature et couvrent l'entièreté de la zone de l'orthoimage.

Le paquetage « *Buildings* » contient plusieurs types de classes. La classe maitresse du paquetage, celle qui contient les géométries de tous les bâtiments de la ville, est *habitatcub0_0207*. La plupart des autres classes contiennent des points dits « d'affectation » proches des bâtiments et donnant des informations sur l'affectation de ceux-ci. Il y a parmi ces classes de nombreuses redondances comme en témoigne notamment la similitude de certains de leurs noms. Certaines classes contiennent également des buildings, les mêmes que dans la classe maitresse, mais en moindre nombre, et qui se situent dans des zones-tampon autour des points d'affectation. La classe *hotel de ville* est entièrement vide tant au niveau géométrique qu'au niveau attributaire. Nous avons également ajouté dans ce paquetage la classe *population figures*. Il s'agit d'un semis de points contenant des informations sur la population vivant en ces lieux en 1987, 1992 et 2005, calculées sur base du recensement de 1976 et de taux d'accroissement[7]. L'intérêt de conserver cette classe au sein du paquetage « *Buildings* » vient du fait que les informations qu'elle contient pourraient être rattachées aux bâtiments. Ce qui permettrait, dans le cadre d'une analyse des pertes humaines potentielles lors de glissements de terrain ou d'inondations, de connaître le nombre de personnes occupant chaque bâtiment.

Les classes du paquetage « *Hydrographie* » reprennent le chevelu hydrographique. Certaines de celles-ci ne s'étendent que sur la zone urbaine de Bamenda tandis que d'autres couvrent une zone bien plus vaste, comme la couche *hydro_satellite towns* qui contient un réseau s'étendant jusqu'à 800km du centre-ville de Bamenda !

De même, dans les classes de relief, les étendues couvertes varient de la seule zone urbaine à l'étalement jusqu'aux villes voisines. Les classes se différencient par l'équidistance entre courbes, équidistance qui est soit de 5m, soit de 50m!

Le paquetage « *Voiries* » reprend toutes les infrastructures accessibles aux véhicules motorisés ou non. Il contient quelques classes relatives au réseau routier ainsi que d'autres couvrant des zones-tampon autour de ses principaux axes. Deux classes de parkings sont présentes sous forme ponctuelle dont l'une, si l'on en croit sa dénomination, semble «*révisée*». Une classe contenant un inventaire des ponts de la ville est également disponible. Enfin, une classe ponctuelle renseignant sur les endroits critiques et à risque des routes est reprise dans le paquetage (*critical areas and trouble spots*).

D'une manière générale, un grand nombre de classes est, en soi, un indice de qualité. De nombreux domaines sont couverts et nous pouvons dire, en première approximation, que tous les ingrédients sont réunis pour pouvoir développer un prototype. L'analyse des données, au chapitre 4, détaillera ce point.

Les traitements de conformité et la rédaction de la documentation ad hoc ont dû être réalisés manuellement car aucun des logiciels retenus ne dispose d'un AGL.

7 La source de ce taux n'a pu être déterminée (NYAMBOD 2010).

3.3 Hypothèse de travail

L’hypothèse forte sur laquelle repose le travail de NYAMBOD, est qu’*un SIG est utile pour gérer les risques naturels de la ville de Bamenda* identifiés par lui, et ce, sous la *contrainte de faisabilité technique.*

Cette hypothèse ne pourra être validée qu’à l’issue de son travail en septembre 2010, après utilisation de notre prototype SIG pour réaliser un modèle de risque.

L’étude de faisabilité technique nous étant assignée, l’hypothèse sur laquelle repose notre propre travail est qu’*avec les données existantes modélisées suivant la norme ISO 19110 et les logiciels Open Source disponibles, il est possible d’implémenter un prototype SIG, utile à gérer certains des problèmes urbains de la ville de Bamenda tels qu’identifiés par l’enquête de NYAMBOD*, et ce, sous la *contrainte de minimisation des coûts*.

Les développements pratiques exposés dans les chapitres qui suivent montrent quelle a été la méthodologie que nous avons suivie, en tant que concepteur, en vue de valider cette hypothèse et de construire un outil opérationnel.

4 Analyse et préparation des données

4.1 Introduction

Selon CALKINS, l'étape qui suit directement l'acquisition des données est la phase d'évaluation de la qualité des données brutes (CALKINS 2001). Elle est à réaliser au niveau des paquetages qui ont été créés au chapitre précédent et porte sur les aspects géométrique et attributaire. Les résultats de cette analyse seront exposés dans un rapport de qualité des données brutes.

Une sélection des données utiles est ensuite opérée sur base de ce rapport. Le but est de ne garder que les données les plus pertinentes et allant dans le sens recherché. Cette tâche a été réalisée en collaboration avec NYAMBOD pour respecter une philosophie de conception de départ selon laquelle le client participerait à l'élaboration de son futur outil de travail (LAURINI 2001).

Afin de les optimiser et préalablement à leur modélisation, les données conservées font l'objet du pré-traitement suivant: les données utiles les plus adéquates sont compilées dans une même classe et lorsqu'il y a une sélection à opérer entre informations similaires, ce sont les plus récentes et les plus précises qui sont conservées. Ces opérations peuvent intervenir à l’intérieur de classes d'un même paquetage uniquement. Cette manière de procéder réduit considérablement le temps d'analyse des traitements d'optimisation entre classes mais oblige le concepteur à former judicieusement les paquetages.

Au niveau géométrique, les traitements se résument principalement à corriger certaines règles élémentaires et naturelles de topologie et à faire des corrections de géométries lorsque cela est possible. Ces corrections sont notamment nécessaires lorsque les formes des objets digitalisés ne correspondent pas parfaitement aux formes des objets correspondants reconnaissables sur l’orthophotographie. Lorsque les défauts sont trop nombreux, la quantité de travail pour les corriger devient trop importante. Les données sont alors laissées inchangées et les défauts sont mentionnés dans le rapport final de qualité des données opérationnelles.

Au niveau attributaire, il faut veiller à ne garder que les informations utiles et supprimer celles qui n'apportent aucune information pertinente ou sont redondantes, ceci dans un souci d’économie d'espace mémoire sur le disque dur et afin de réduire le temps de requête lorsque la base sera opérationnelle. Il est aussi important d'attribuer des noms de champs uniques au sein des classes gardées afin d'éviter de possibles confusions dans l'identification des champs lors de requêtes sur la future base de données.

L'ensemble des classes opérationnelles, c'est-à-dire sélectionnées, structurées et optimisées, est intégré dans un catalogue créé suivant la norme ISO 19110, dans le but de :

- présenter, aux futurs utilisateurs du prototype, une description claire, précise et sans ambiguïté de l'ensemble des données opérationnelles ;
- faciliter les tâches de mises à jour des données ;
- les rendre interopérables et donc utilisables par les futurs concepteurs qui souhaiteraient extrapoler le prototype vers un SIG complet pour la ville de Bamenda.

4.2 Analyse de la qualité des données brutes

4.2.1 Orthophotographie

L'orthophotographie, représente une vue aérienne de la zone basée sur le centre-ville de Bamenda et est en format *tiff*. Il est accompagné d'un fichier *World File for Tiff* (tfw) qui contient les coordonnées X et Y, en mètres, du coin supérieur gauche de l'image dans le plan de projection, la largeur et la hauteur en mètres des pixels et les paramètres de rotation sur les lignes et sur les colonnes de l'image.

Le référentiel spatial dans lequel sont exprimées ces coordonnées nous a été communiqué par NYAMBOD et est repris dans la figure 8.

Figure 8. Paramètres géodésiques et de projection

Pays	Système géodésique	Ellipsoïde associé	Projection	Zone
Cameroun	WGS84	GRS80	UTM	32 N

(Source : NYAMBOD 2010)

Pour l'image correspondante, *bamenda.tif*, les paramètres sont les suivants :

Dimension X du pixel: 1.00
Valeur du paramètre de rotation sur les lignes : 0.00
Valeur du paramètre de rotation sur les colonnes : 0.00
Dimension Y du pixel: -1.00
Abscisse du coin supérieur gauche : 623350,66
Ordonnée du coin supérieur gauche : 663204,58

Il faut remarquer que la valeur négative de la dimension Y du pixel vient du fait que, dans l'image raster, l'axe est dirigé du haut vers le bas.

Les valeurs nulles des paramètres de rotation sur les lignes et les colonnes signifient que le nord de l'image correspond au nord géographique.

La zone couverte peut facilement être déterminée à partir de ces indications et de la connaissance du nombre de lignes et de colonnes qui composent le document. Ces dernières informations sont stockées dans les métadonnées du fichier :

Nombre de lignes : 10.072
Nombre de colonnes : 9.584

Étant donné que la résolution géométrique est de 1m tant en X qu'en Y, la zone couverte par l'orthophotographie est un rectangle de 10,072 km sur 9,584 km dont le coin supérieur gauche a pour coordonnées (623350,66 663204,58) dans le plan de projection associé.

L'orthophotographie date de février 2006 et est issue d'une orthorectification d'une image aérienne brute enregistrée par les capteurs du satellite *Ikonos*. La connaissance de la saison au cours de laquelle la prise de vue a été faite est d'une importance capitale car les phénomènes saisonniers caractérisent le paysage et le modifient. L'aspect, voire la forme des objets varie en fonction de ces phénomènes temporels Dans le cas présent, les objets digitalisés sur base du document aérien sont caractéristiques de la saison d'hiver au Cameroun. Les conditions climatiques aux conséquences locales perceptibles ayant existé à ce moment là ainsi que durant les semaines et jours précédents ont également un impact sur le contenu du document.

Même le moment de la journée a une influence sur la perception des objets. La luminosité du sol et des objets dépend de l'heure de prise de vue (heure impossible à déterminer dans le cas présent) et a un impact sur le contraste de l'image ainsi que sur la perception des teintes. L'ombre portée par les objets hauts, d'autant plus importante que le moment de la prise de vue s'éloigne du midi local, peut masquer des objets plus petits. Tout ceci influence voire perturbe la personne qui digitalise et des erreurs de digitalisation peuvent être commises.

Les variations de forme et d'aspect des divers objets reconnaissables sur l'orthophotographie peuvent compromettre le caractère de référence du document et augmenter le risque de confusions. En voici quelques exemples :

- variation de hauteur des plans d'eau en fonction de la saison des pluies et de la saison sèche ;
- extension des zones humides variant également en fonction de la saison ;
- variation de la luminosité des sols pouvant fausser l'interprétation de leur affectation ;
- altération de la géométrie des objets par des phénomènes de courte durée tels que des coulées de boue et des ruptures de berges provoquant un étalement des zones humides qui pourrait laisser croire à une humidité permanente.

Il faudra intégrer tous les aspects cités ci-dessus et en tenir compte dans les analyses de données.

4.2.2 Les données Shapefile

Avant même d'ouvrir quelque fichier que ce soit, il faut entreprendre l'analyse du dossier contenant les fichiers car il fournit déjà de nombreuses informations, surtout en ce qui concerne la complétude des métadonnées.

Un *Shapefile* est composé d'un ensemble de fichiers de mêmes noms mais d'extensions différentes. Des fichiers connexes peuvent être fournis avec les *Shapefile* et nous les citons ci-dessous (liste non-exhaustive) ; cependant seuls les trois premiers sont obligatoires :

- ⇨ *Main file* (**.shp*) : stockant la géométrie des entités géographiques ;
- ⇨ *Data Base file* (*.dbf) : contenant les données attributaires ;
- ⇨ *Index File* (*.shx) : index reprenant les pointeurs vers les différentes entités de *Main file ;*
- ⇨ *Projection File* (*.prj) : contenant les paramètres du datum et de la projection, dans lesquels s'insèrent les données.

Dans le cas des données reçues de l'Administration principale de Bamenda-ville, tous les *Shapefile* contiennent systématiquement trois fichiers dont les suffixes sont *.shp, shx* et *.dbf*. Pour certains d'entre eux, des fichiers *.sbn* et *sbx* sont également disponibles, mais aucun n'est accompagné d'un fichier *.prj* ! Ce dernier fichier n'est pas indispensable à la visualisation, mais il l'est à la description des métadonnées et à l'exportation des données. Il faut en effet connaître les paramètres de projection afin d'effectuer les transformations de coordonnées nécessaires pour l'intégration dans, par exemple, un modèle de risque. Nous pouvons cependant supposer que le référentiel est le même que celui de l'orthoimage car les *shp* ont été digitalisés sur base de celui-ci. Rappelons qu'il s'agit de la projection UTM zone 32 N à partir du datum WGS84.

4.2.2.1 Qualité temporelle

Les données n'ont pas été digitalisées à la même période que la prise de vue de l'image aérienne. Elles l'ont été en juillet 2007, soit un an plus tard. Certaines d'entre-elles n'étant plus d'actualité, NYAMBOD les a revues et mises à jour en juillet 2009, sur base de sa connaissance de terrain. Les mises à jour portaient essentiellement sur les fonctions de bâtiments et les parkings ainsi que sur certaines données administratives qui avaient été modifiées, au niveau du Cameroun, à la fin de l'année 2007 et donc, après digitalisation. Rappelons que les fichiers retouchés portent le suffixe 0707.

4.2.2.2 Exactitude de position et précision de digitalisation

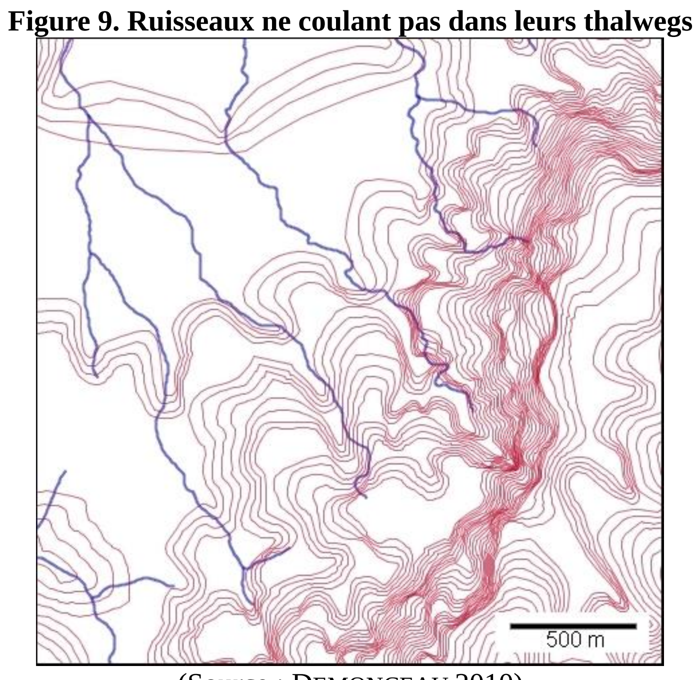
La qualité du positionnement des objets géographiques dans le plan de projection ainsi que la qualité du respect des contours des géométries sont évaluées par la superposition de couches d'objets sur l'orthophotographie considérée comme correctement positionnée dans le plan de visualisation. Il n'y a pas d'autres documents disponibles pour juger de la qualité du positionnement des objets géographiques dans l'espace. Seules certaines données ont une emprise au sol bien visible et identifiable (routes, rivières, bâtiments, etc.). L'analyse ne pourra porter que sur celles-ci et non sur d'autres données telles que frontières administratives, carrefours critiques, points d'affectation des bâtiments, etc.

4.2.2.2.1 Courbes de niveau

Les couches contenant les courbes de niveau semblent toutes avoir subi une translation d'une soixantaine de mètres vers l'ouest par rapport à leur position sur l'orthophotographie. En effet, bien qu'il ne soit pas possible d'identifier les courbes de niveau sur l'orthophotographie, deux éléments permettent de confirmer cette hypothèse :

- le décalage entre le tracé des cours d'eau et la représentation des thalwegs par les courbes de niveau, lorsque la couche d'hydrographie est superposée aux couches contenant les courbes de niveau, comme illustré dans la figure 9;

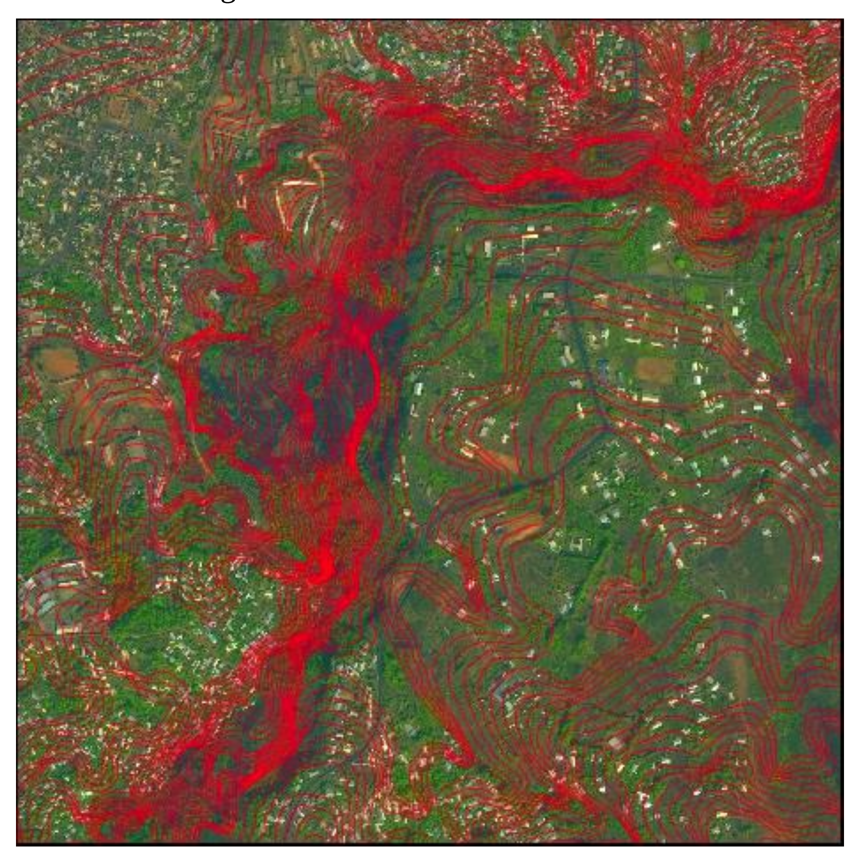
Figure 9. Ruisseaux ne coulant pas dans leurs thalwegs



(Source : DEMONCEAU 2010)

– la translation des courbes de niveau, par comparaison entre la couche d’hydrographie et le relief perceptible sur l'orthophotographie, comme illustré à la figure 10;

Figure 10. Courbes de niveaux translatés



(Source : DEMONCEAU 2010)

Plusieurs raisons ont pu provoquer ce décalage :

- une différence de système de projection entre les couches ;
- des erreurs apparues lors de l'orthorectification.

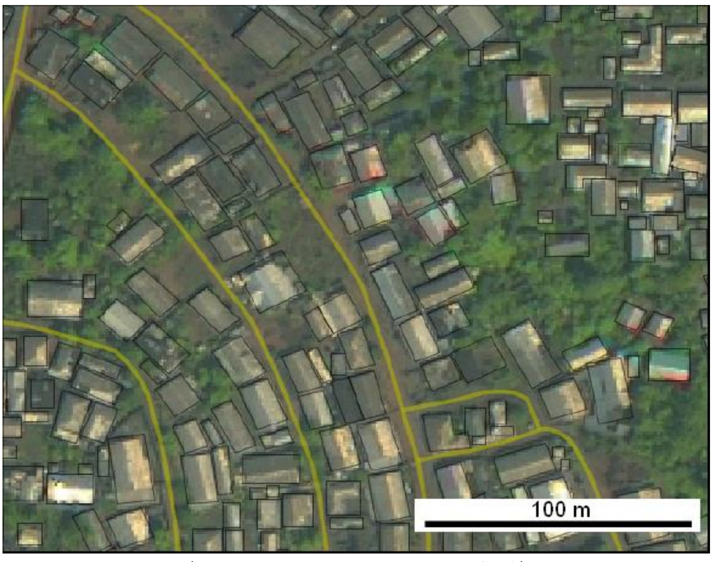
4.2.2.2.2 Couches administratives

Suite aux multiples discussions menées avec NYAMBOD à ce sujet, il ressort que seule la couche S*ubdivibdapolygon0_0707* représentant les frontières des arrondissements du département dans lequel se situe la ville de Bamenda est correcte. Les autres couches polygonales sont à rejeter car, d'après lui, elles ne sont nullement caractéristiques des frontières réelles. Il est important de noter que des couches ponctuelles correspondant aux centres des quartiers et des villages sont également disponibles et peuvent potentiellement faire partie des données opérationnelles via des traitements adéquats les valorisant. Cet aspect sera discuté dans la section 5.2.1.1. relative à la création des classes.

4.2.2.2.3 Autres classes

La qualité des géométries des objets présents dans les autres classes est satisfaisante. Les contours des polygones et le tracé des objets linéaires respectent relativement bien ceux des objets homologues visibles sur l'orthophotographie. C'est ce qu'illustre la figure suivante dans le cas de deux couches choisies au hasard, *Roads0_0707* et *Habitatcub0_0207*, pour une zone de la ville également choisie au hasard.

Figure 11. Qualité de la digitalisation des couches de routes et de bâtiments avec l'orthophotographie en fond



(Source : DEMONCEAU 2010)

Il faut savoir que l'exactitude de position ainsi que la précision de la digitalisation sont intimement liées à la référence temporelle. Nombreux sont les objets qui, au cours du temps, peuvent changer de forme (bâtiments, zones humides, routes, etc.) ou de place (fonctions de bâtiments, zones routières critiques, etc.). NYAMBOD a cependant effectué une mise à jour pour certaines des couches relatives à ces objets (voir sous-section 4.2.2.1).

4.2.2.3 Qualité de la complétude

Certaines classes ne sont pas exhaustives c'est-à-dire qu'elles ne contiennent pas l'entièreté des objets présents et visibles sur le document aérien. Cette problématique est révélée par une première analyse visuelle, ne portant que sur des objets bien visibles et identifiables, du résultat de la superposition entre les couches et l'orthophotographie.

Le manque de complétude peut également être lié à la référence temporelle choisie. Dans notre cas, il y a un an d’intervalle entre l'acquisition de l'orthophotographie et la digitalisation des couches. Des objets peuvent apparaître et d'autres disparaître durant ce laps de temps. N'oublions pas que NYAMBOD a retouché les couches relatives à ces objets sur base de sa connaissance personnelle de terrain.

Une fois les défauts détectés, les outils d'analyse spatiale sont utilisés pour évaluer le taux de complétude dans le relevé des objets.

4.2.2.3.1 Paquetage Routes – Ponts (*Bridges and Culverts*)

La présence d'un pont est théoriquement justifiée par le croisement, à des altitudes différentes:

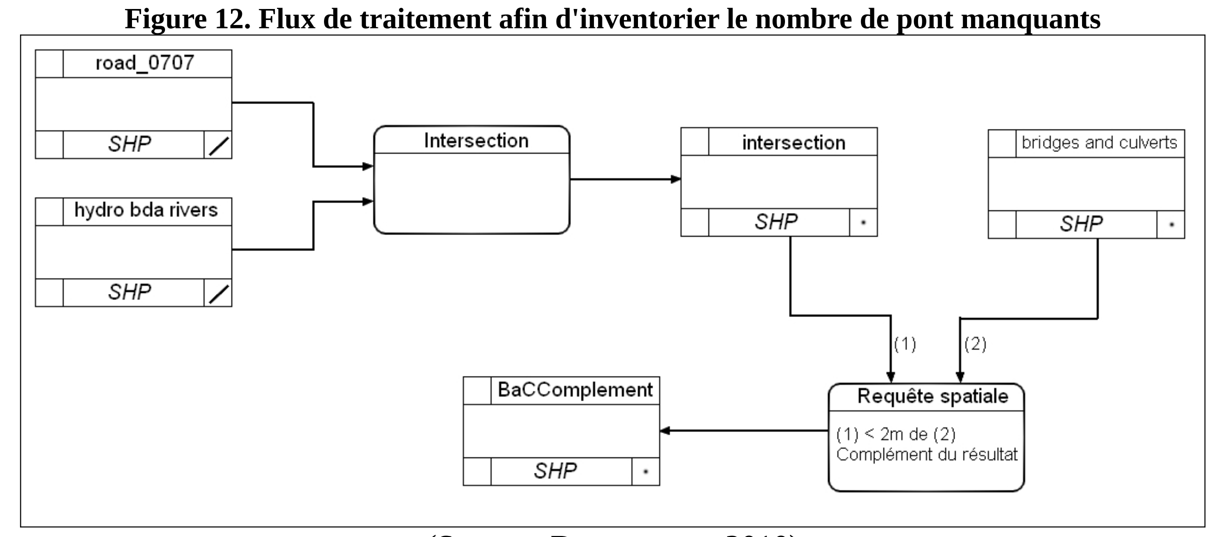
- d’une route et d’un cours d'eau (ruisseau, fleuve ou caniveau encore appelé dalot)
- de deux routes
- de deux routes et d’un cours d'eau (triple croisement et forcément à trois niveaux).

Cependant, il s'avère que des ponts ne sont pas présents bien qu’une de ces conditions soit remplie. Plusieurs opérations sont réalisées afin de quantifier ce manque de complétude.

La couche contenant le réseau routier est intersectée avec la couche contenant le réseau hydrographique. Le résultat en est une couche contenant les points d'intersection. De ces points, il faut retirer les ponts déjà présents dans la couche de ponts d'origine. La couche finale est comparée avec l'orthophotographie pour une analyse visuelle de la nature de ces points.

La figure suivante illustre un exemple de flux de traitements permettant d'offrir une valeur chiffrée du nombre d'« oublis » et de la comparer avec le nombre d'objets présents dans la couche des ponts.

Figure 12. Flux de traitement afin d'inventorier le nombre de pont manquants



(Source : DEMONCEAU 2010)

L'analyse de la couche associée à la classe *BaCComplement* indique que 50 points résultant des intersections ne correspondent pas à des entités présentes dans la couche des ponts d'origine. Ces intersections complémentaires sont analysées sur l'orthophotographie et il en ressort que :

- pour 29 points, la nature ne peut pas être déterminée à cause de problèmes de visibilité (présence de végétation obstruant la vue, zone peu claire …) ;
- pour 19 points, il s'agit bel et bien de ponts clairement identifiables ;
- pour 2 localisations, il s'agit clairement de passages à gué sans aucune infrastructure de traversée.

Il est important de spécifier que cette analyse est uniquement réalisée sur base visuelle et présente donc un certain degré de subjectivité. De plus, il est possible que de nouveaux ponts aient été construits et que d'autres aient été détruits entre le moment d'acquisition de l'image aérienne et celui de la digitalisation. Dès lors, étant donné qu'une vérification sur terrain n'a pas été entreprise et que la mise à jour faite n'a pas concerné la couche de ponts, nous avons attribué à ces cinquante points supplémentaires la valeur « *To determine* » (= à déterminer) pour l'attribut *BAC_TYPE* (= type de pont).

Par contre, si l'on considère que les 19 ponts non-inventoriés dans la couche des ponts constituent bel et bien des oublis de la part des personnes responsables de la digitalisation, le taux de complétude pour ce type de données est donc de 82,57%, ce qui peut être considéré comme satisfaisant si les principaux ponts de la ville sont relevés.

Un pont peut également exister lorsqu'un chemin piétonnier passe au-dessus d'une route, d'une rivière, voire d'une vallée sèche. Cependant, dans ces cas particuliers, le manque de complétude ne pourra être quantifié vu la non-disponibilité des couches requises.

4.2.2.3.2 Paquetage Bâtiments – Bâtiments (*Habitatcub0_0207*)

Le fichier contenant la couche des bâtiments est le plus volumineux de tous les *shp*. Le respect de certaines règles fondamentales de topologie directement liées à des entités de type « bâtiment » est vérifié à ce stade de l'analyse :

- un bâtiment est uniquement représenté par une primitive polygone simple (pas de multi-géométrie, ni entité de type ligne ou point) ;
- les polygones doivent être correctement fermés ;
- les polygones ne doivent pas s'auto-intersecter ;
- deux bâtiments différents ne peuvent pas s'intersecter (ce qui implique aussi qu'un polygone ne peut pas en contenir un autre plus petit) :
- un bâtiment est représenté par un unique polygone (pas de doublons géométriques).

Cette vérification a donné les statistiques suivantes :

Nombre de polygones présents dans la couche : 45.016
Nombre de polygones fermés : 44.998
Nombre de multi-géométries : 32
Nombre de points présents (géométrie inadaptée) : 17
Nombre de polygones auto-intersectants : 1
Nombre de bâtiments qui s'intersectent : 2.432
Nombre de bâtiments complètement inclus dans un autre : 12
Nombre de doublons géométriques : 1.775
Nombre de bâtiments contigus : 1.590

Avant de pouvoir intégrer cette classe dans la base, un sérieux nettoyage au niveau topologique doit être entrepris.

Les multi-géométries peuvent causer des problèmes lors du chargement dans certains logiciels SIG. C'est le cas pour ceux n'acceptant que des objets sous forme de primitive géométrique simple à savoir le point, la ligne et le polygone. Des problèmes peuvent également être liés au SGBD dans lequel les données sont chargées ; par exemple, *PostGIS* n'accepte qu'un type de géométrie par couche. La présence de ces géométries inadéquates est sûrement due à une distraction de la part de la ou des personne(s) ayant digitalisé la couche d'objets.

L’existence de doublons de géométries de bâtiments ne cause pas de problème au niveau de l'exploitation à partir d'un SIG-logiciel ou d’un SGBD, mais elle est désagréable d'un point de vue topologique et elle fausse les éventuels calculs statistiques. En outre, la mise à jour dans la future base de données de la géométrie de tout bâtiment concerné par un doublon doit être réalisée deux fois au lieu d'une car cela conduit à des incohérences au sein de la base. De plus, la présence de doublons occupe inutilement de l'espace disque et allonge le temps des requêtes lorsqu'elles ont trait à cette couche. Il est donc préférable de les supprimer.

Même si, au premier abord, la couche semblait être complète, une analyse plus approfondie a fait ressortir l’absence de quelques entités. C'est le cas pour certains bâtiments se trouvant près des points d'affectation de buildings (voir figure 13). Sept bâtiments ont été rajoutés par digitalisation manuelle, ce qui donne un taux de complétude de 99,98 % pour ce type de données et est un excellent résultat.

Figure 13. Bâtiments non-digitalisés près de points *Functions*

(Source : DEMONCEAU 2010)

4.2.2.4 Qualité sémantique

L'analyse de la sémantique porte sur les données attributaires des tables *.dbf* associées aux *main file shp*. Elle porte sur la précision et l'exactitude des attributs. En d'autres mots, elle permet de vérifier que la description des données est en accord avec la réalité.

Notre manque de connaissance du terrain et l'absence de documents fiables ne nous permettant pas d'évaluer la fiabilité de ces attributs, cette analyse de la sémantique n'a pu être réalisée.

La description exhaustive de l'ensemble des classes se trouve en annexe, et pour chaque classe, les attributs, les types de données et l'occurrence de la valeur de chaque attribut dans le *dbf*, sont décrits.

4.2.2.5 Cohérence logique des données

L'analyse des tables *dbf* fait tout d'abord ressortir un problème de langue. L'anglais et le français s'y côtoient aléatoirement, alors que pour éviter toute ambiguïté et tout risque de confusion (notamment via ce qui est communément appelé les *false friends*), il serait préférable d'utiliser une seule langue.

Une analyse des noms de classes a été réalisée afin de quantifier respectivement le nombre de noms en anglais, en français et en « *franglais* ». Elle a donné les résultats suivants :

Noms de classes en français : 15
Noms de classes en anglais : 58
Noms de classes en *franglais* : 2

Il faut remarquer que la somme de ces trois nombres ne correspond pas au nombre de *shp*. Certains noms (*finance*, *education*, *evolution*,...) ont la même orthographe dans les deux langues et ont donc été comptés deux fois, une fois dans les noms en français et une fois dans les noms en anglais.

Que ce soit en anglais ou en français, nombreuses sont les fautes de frappe et d'orthographe dans les noms de classes, de champs et de valeurs d'attributs. En voici quelques exemples :

- *« Artirary road buffer »* au lieu de « *Artiliary road buffer* » (nom de classe)
- *« Gares routiére »* pour l'attribut « *Code_nom* » de la classe « *Eauip machands* »
- *« Organasitions internationale »* pour l'attribut « *Nom_code* » dans la classe « *Equipe public et parapublic* »

L'ensemble de ces fautes peut être détecté à l'aide du dictionnaire intégré au logiciel utilisé pour visualiser les données attributaires (CALKINS 1997). Vu le grand nombre de noms propres imbriqués dans les tableaux, la correction aurait dû être menée en passant en revue, un par un, chaque mot détecté par le dictionnaire. Le nombre de mots détectés est cependant trop élevé pour se lancer dans l'analyse individuelle de chaque mot ; c'est pourquoi les corrections sont faites uniquement au fur et à mesure de la détection des erreurs.

Autre erreur ayant des répercussions moins graves à ce stade de l'analyse, les noms de champs ne sont pas toujours uniques au sein des classes. Certains champs portent les mêmes noms mais ne font pas référence aux mêmes informations, tandis que d'autres le font mais portent des noms différents. Ce type d'erreur aura un impact négatif lors de l'exploitation de la base de données dans l'identification des champs désignés pour les requêtes. Les données feront également l'objet d'un inventaire au sein d'un catalogue normalisé ISO 19110 qui proscrit l'utilisation d'un même nom pour plusieurs champs au sein de l'entièreté de la base.

De nombreux noms contiennent des mots imaginaires et absents du dictionnaire et leur interprétation est laissée à l'imagination du lecteur.

La figure 14 reprend l'ensemble des champs dont la signification n'a pu être déterminée (mots imaginaires ou non) :

Figure 14. Inventaire complet des noms de champs non identifiés

Classe	*Attribut*	*Justification*
Landuse	Code	Cet attribut pourrait se rapporter aux classes d'ulisation de sol mais ce n'est pas le cas car il n'et pas possible de trouver une logique de classification sur base de ces codes
	Sub_Code	Idem
Landuse_study_area	Code	Idem
	Sub_Code	Idem
Parking facilities revised	Pos_situti	D'autant plus difficile à déterminer étant donné que tous les champs sont vides
Road network	Priorite	Toues les valeurs sont à zéro
	Action_pro	Aucune valeur donnée pour cet attribut
	Départ	Idem
Roads_0770	Priorité	Idem que pour l'attribut portant le même nom de la classe Road network
Service public et parapublic	Status	Aucune logique géographique ou par rapport aux autres attributs
Landuse points1207	Lu	Les acronymes faisant partie des valeurs que peuvent prendre les attributs n'ont pas pu être identifié (et ils ne correspondent même pas à une classe ou sous-classe d'utilisation du sol)

(Source : DEMONCEAU 2010)

Il y a aussi des colonnes-doublons au sein des mêmes tables c'est-à-dire que tant le noms des champs que les valeurs pour chaque entrée sont identiques. Ces doublons sont à éliminer tout comme les champs qui ne sont d'aucune utilité (champs vides, acronymes non-identifiables, champs résultant d'un calcul réalisé sur d'autres champs,…).

La totalité des 70 fichiers *dbf* associés aux *shp* a été soumis à l'examen des fautes d'orthographe et des mots imaginaires des nom de champs.

4.3 Rapport qualité

Suite à l'analyse de la qualité aux niveaux sémantique et géométrique, nous pouvons conclure que pour être opérationnelles, les données nécessitent une mise en forme. Celle-ci ne sera appliquée que sur les données effectivement conservées pour faire partie du prototype et consistera en :

- un ensemble de corrections des données tabulaires telles que suppression de champs redondants identiques, renommage … ;
- des corrections géométriques pour respecter les topologies ;
- un calage géométrique de la couche des courbes de niveau.

D'une façon générale, les données acquises sont de piètre qualité, tant au niveau géométrique qu'au niveau attributaire. La plus importante lacune concerne l'absence de tout fichier *prj* pour les *shp*. Il est toutefois possible d'y remédier en connaissant le datum et la projection qui ont été employés étant donné que les *shp* ont été digitalisés sur l'image dont le référentiel est connu.

Les données reprises dans le paquetage « *Administratif* » sont de qualité très médiocre. Seule une couche polygonale a pu être retenue (*subdivibdapolygon0_0707*), les autres devront faire l'objet d'un pré-traitement. De même, il n'a pas été possible de vérifier l'exactitude des limites administratives et nous avons dû faire entièrement confiance à NYAMBOD pour la sélection des données utiles à conserver dans le paquetage « *Administratif* ».

La qualité sémantique n'est guère meilleure (problèmes de langue, nombreux champs à éliminer, ..).

Malgré tous ces éléments négatifs, il faut reconnaître certaines qualités à ces données : la mise à disposition d'une orthophotographie à 1m de résolution couvrant entièrement la zone urbaine de la ville ; la quasi exhaustivité des couches de routes, de cours d'eau et de bâtiments ainsi que, dans une moindre mesure, de la couche de ponts.

La section relative à la construction des classes aura pour but de valoriser les données et d'en exploiter un maximum d'informations.

4.4 Sélection de données utiles

La tâche présente est de sélectionner les données utiles pour le prototype à concevoir. La méthodologie adoptée consiste à éliminer progressivement les données ne rencontrant pas les objectifs fixés.

Les données présentant les conditions suivantes, classées par ordre décroissant d'importance, seront éliminées :
- être vides, c'est-à-dire ne contenir aucune géométrie ni attribut ;
- être relatives à un thème qui n'a aucun rapport avec l'objectif de l'application ;
- présenter des contours de géométrie ne correspondant à aucune limite réelle ;
- correspondre au résultat d'un traitement géométrique sur une ou plusieurs autres classes[8] ;
- consister en une redondance d'informations.

Les données rejetées ainsi que les causes de rejet, sont résumées à la figure suivante (par ordre alphabétique) :

8 L'intégration, dans le prototype, de données respectant cette condition, ne ferait qu'ajouter des données qui pourraient ultérieurement être obtenues via requêtes effectuées sur données corrigées et validées.

Figure 15. Ensemble des classes rejetées et la cause des rejets

Classe rejetée	***Cause***	***Identifiant de classe***
artirary road buffer	Classe issue d'une opération géométrique sur la classe roads_0707	1
buffer round critical junctions	Classe issue d'une opération géométrique sur la classe critical areas and trouble spots	4
catchment sec general and technical	Classe issue d'une opération géométrique	5
cbd	Classe issue d'une opération géométrique	6
evolution	Classe dont la géométrie des polygones ne correspond à aucun découpage administratif existant	18
evolving hierarchy of settlements	Classe ne contenant aucune donnée pertinente	19
gradient analysis	Classe dont la géométrie des polygones n'a pas pu être déterminée	21
hotel de ville	Classe vide et donc inutile	24
houses whitin catchment zone	Classe issue d'une opération géométrique entre habitatcub0_0207 et primary catchment zone	25
line covering sat towns	Classe issue d'une opération géométrique	41
nursery school travelling distance govt	Classe issue d'une opération géométrique sur l'objet nursery de la classe education0_0707	42
primary catchment zone	Classe issue d'une opération géométrique	46
primary road buffer	Classe issue d'une opération géométrique sur la classe roads_0707	47
quartier polygon	Classe dont les géométries des quartiers ne correpondent à aucun découpage administratif	52
road accessibility	Classe issue d'une opération géométrique sur la classe roads_0707	55
secondary road buffer	Classe issue d'une opération géométrique sur la classe roads_0707	59
subregional context_merge	Classe ne contenant aucune information pertinente	66
tertaty road buffer	Classe issue d'une opération géométrique sur la classe roads_0707	69
travelling distance sec school	Classe issue d'une opréation géométrique sur la classe education0_0707	70

(Source : DEMONCEAU 2010)

4.5 Construction des classes

Le but de cette nouvelle étape est d'éliminer les classes restantes faisant encore l'objet d'informations redondantes au niveau géométrique ou attributaire. Les informations les plus pertinentes au niveau des objets sont ainsi compilées au sein d'une seule et même classe afin de créer des classes les plus complètes possible au niveau information.

Les classes finales créées sont celles qui, d'un point de vue géométrique :

- sont les plus précises (en comparaison avec l'orthophotographie) ;
- sont les plus exactes ;
- sont les plus complètes.

Ces classes sont complétées au niveau attributaire, si nécessaire, lorsque des informations externes utiles et pertinentes sont disponibles dans d'autres classes du même paquetage.

Un attribut clé sera ajouté à chacune des classes d'objets. Le but étant de pouvoir identifier, de manière non-équivoque, un objet au sein de chacune des classes.

En résumé, les classes construites sont celles dont les géométries sont les plus adéquates, en rapport avec l'orthophotographie, et dont les tables d'attributs sont les plus complètes.
Les classes finales sont les suivantes :

Figure 16. Classes potentiellement utilisables

Paquetage	***Classes***	Primitive géom.	Identifiant de classe
Administratif	bamenda satellite towns	Point	2
	internal urban spatial structure	Polygone	30
	limit urban area bamenda	Ligne	38
	limit urban area bamenda line	Ligne	39
	limite arrondissement	Polygone	40
	quarters mendankwe	Point	48
	quarters_bamenda	Point	49
	quarters_mendankwe	Point	50
	quarters0_0707	Point	51
	subdivibdapolygon0_0707	Polygone	63
	Subdivibdapolyline0_0707	Ligne	64
	subdivision line	Ligne	65
Affectation de sol	hydro bda lakes	Polygone	26
	landuse	Polygone	32
	landuse points	Polygone	33
	landuse points1207	Point	34
	landuse polygons 1207	Polygone	35
	landuse_study area	Polygone	36
	swamps	Polygone	67
Buildings	council halls	Polygone	8
	eauip marchands	Point	12
	education	Point	13
	education0_0707	Point	14
	equipe public et parapublic	Point	15
	equipements d'hôtellerie et touristiques	Point	16
	equipements marchandes	Point	17
	finance	Point	20
	habitatcub0_0207	Polygone	22
	health	Point	23
	industry	Point	29
	land marks	Point	31
	population figures	Point	45
	religious institution	Point	53
	securité publique	Point	60
	service public et parapublic	Point	61
	télécommunication	Point	68
Hydrographie	hydro bda rivers	Ligne	27
	hydro_satellite towns	Ligne	28
	limit satillite town area	Ligne	37
	rivers sat towns	Ligne	54
Relief	contour satellite towns	Ligne	7
	courbes de niveau bamenda	Ligne	9
	courbes de niveau bamenda_urban area	Ligne	10
	station escarpment	Ligne	62
Voiries	bridges and culverts	Point	3
	Critical areas and trouble spots	Point	11
	parking facilities revised	Point	44
	road network	Ligne	56
	roads linkin the evolving hieracchy of satellite towns	Ligne	57
	Roads_0707	Ligne	58

(Source : DEMONCEAU 2010)

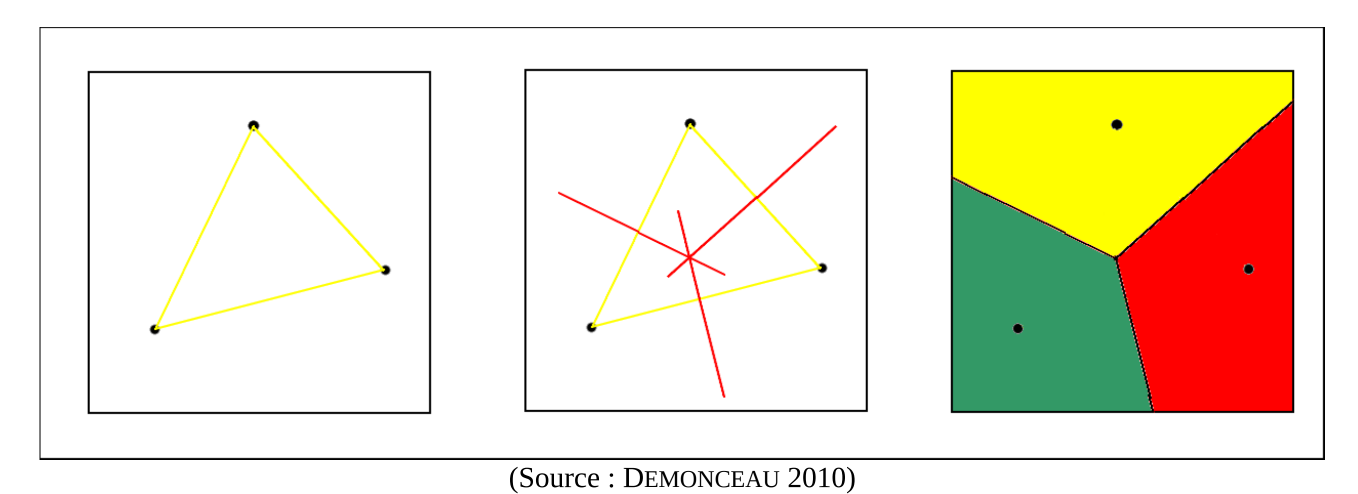
4.5.1 Paquetage « Administratif »

L'analyse des données brutes a révélé que les données relatives aux arrondissements de la ville, *Subdivibdapolygon0_0707*, étaient a priori les seules pertinentes parmi les classes du paquetage administratif. Elles sont donc conservées.

La couche *quarters0_0707* est une couche ponctuelle, ce qui ne constitue pas la représentation la plus pertinente pour les quartiers. L'idée est de transformer les points de cette couche en polygones, en procédant à une « tessellation » de Thiessen. Cette dernière consiste à créer des zones polygonales centrées sur les entités ponctuelles. Ces polygones sont construits de la façon suivante (figure 17) :

- création d'un réseau de triangles selon le critère de Delaunay ;
- délimitation des contours des polygones par détermination des médiatrices des côtés des triangles

Figure 17. Tesselation de Thiessen



(Source : DEMONCEAU 2010)

Il s'agit d'un cas non-pondéré car aucun attribut n'intervient dans l'élaboration de ces polygones. D'après DONNAY, une tesselation de Thiessen est la manière la plus courante de passer du mode ponctuel au mode zonal s'il n'y a pas de pondération par attributs (DONNAY 2007a). La couche de sortie est appelée *Quarters*.

Le même procédé est appliqué à la couche de points contenant les centres de village. NYAMBOD a cependant opéré une correction des frontières afin qu'elles s'ajustent au mieux à la représentation qu'il a des frontières réelles. La couche *Villages* est ainsi créée.

Enfin, la couche *limit urban area bamenda line*, qui contient une ligne encadrant la zone d'étude, est convertie en polygone de même contour. Ce polygone carré ne représente pas les limites de la ville mais a pour unique vocation de fournir des informations sur la ville en procédant à une jointure spatiale entre les données dans la future base. Le polygone résultant est repris dans la couche *Town*.

4.5.2 Paquetage « Affectation de sol »

Ce paquetage reprend toutes les classes relatives à l'affectation du sol. Cinq d'entre elles couvrent le même territoire que l'orthophotographie tandis que deux autres ne reprennent, sous forme polygonale, que le réservoir et le lac pour l'une (*hydro bda lakes*), et les zones humides pour l'autre (*swamps*).

Une première série d'éliminations de classes peut être immédiatement opérée:

- *landuse points1207* : classe représentant l'affectation du sol sous forme ponctuelle (points éventuellement issus d'un simple sondage aléatoire), ce qui n'est pas une géométrie adéquate pour illustrer l'affectation du sol, et dont la table d'attributs contient des acronymes dont la signification n'a pu être déterminée ;
- *landuse polygons 1207* : classe contenant des polygones de formes variées et dont la position dans l'espace n'a pu être mise en correspondance avec les éléments de terrain. L'aspect attributaire de cette couche ne donne aucune autre information ;
- *landuse points* : classe consistant en une parfaite copie de *landuse* et constituant donc une redondance ;
- *landuse_study area :* classe contenant les mêmes géométries que *landuse* mais dont l'étendue spatiale est plus restreinte ;
- *swamps* : classe qui contient les zones marécageuses sous forme de polygones, mais dont la géométrie est exactement la même que l'objet « *swamps* » dans la couche « *landuse* ».

Les seules classes conservées sont dès lors *hydro bda lakes* et *landuse*. La deuxième contient des objets de type « lac » et « réservoir » mais leur géométrie épouse moins bien les contours identifiés

sur la photographie aérienne que la géométrie des objets homologues dans la classe *hydro bda lakes*. Par exemple, le lac dans *landuse* semble déborder en toutes directions, ce qui peut faire penser à un état de hautes eaux, tandis que dans la classe *hydro bda lakes*, il s'agirait plutôt d'un état de basses eaux (figure 18 et 19).

Quoiqu'il en soit, c'est la géométrie la plus précise par rapport à l'image aérienne, c'est-à-dire celle de *hydro bda lakes* (figure 19), qui sera gardée et transférée dans *landuse*. Cette opération sera réalisée pour respecter la condition, citée précédemment, de ne garder que les géométries les plus précises au sein des classes d'un même paquetage.

Figure 18. Défaut de digitalisation

(Source : DEMONCEAU 2010)

Figure 19. Bonnes géométries à transférer



(Source : DEMONCEAU 2010)

D'une manière pratique, le transfert de géométrie n'est pas simple car « lac » et « réservoir » de *hydro bda lakes* sont contenus dans un seul et même polygone. La première tâche est donc de scinder cette géométrie pour isoler le contour des lacs. Les géométries correspondantes sont ensuite remplacées dans *Lake* et *Dam* de *landuse* par copie de la géométrie en format *Well-Known Text (WKT).*

4.5.3 **Paquetage « Buildings »**

4.5.3.1 Suppression des redondances géométriques

La plupart des redondances d'informations géométriques et attributaires se situent au niveau des couches de classes d'affectation.

La première suppression de redondance concerne les classes relatives à l'éducation : *education* et *education_0707* sont similaires mais seule la dernière est gardée étant donné qu'elle est issue d'un recensement plus récent (Cf le suffixe 0707).

La classe *council halls* est, elle aussi, supprimée puisque les informations qu'elle contient sont déjà présentes dans la classe *service public & parapublic*.

C'est la même raison qui justifie la suppression des classes *eauip marchandes* et *equipe public & parapublic :* des données plus récentes se trouvent respectivement dans les classes *equipement marchandes* et *service public & parapublic* (d'après NYAMBOD).
La classe *land marks* est de nature un peu spéciale, dans le sens où la plupart des objets qu'elle contient sont déjà dans d'autres classes telles *equipement marchandes, health,* ou *religious institution*, pour ne citer qu'eux. Mais une recherche approfondie et minutieuse révèle qu'elle apporte des informations

sur des objets non encore répertoriés dans les autres classes. Il est donc nécessaire d'identifier ces objets et de les ajouter aux classes auxquelles ils doivent appartenir.

Nous devons toutefois rester attentifs car nous ne disposons pas de la date de recensement des objets. Il se pourrait que la classe *land marks* soit plus ancienne que toutes les autres classes et qu'elle contienne des données obsolètes. Dès lors, pour plus de sécurité et pour éviter tout problème d'omission, ces données sont rajoutées aux autres classes correspondantes. Rappelons qu'il s'agit de concevoir un prototype SIG sur des risques naturels, et qu'il est préférable d'avoir trop de données de ce genre sur les bâtiments que trop peu.

Au niveau des polygones de la classe *habitatcub0_0707*, la section dédiée à l'analyse de la qualité des données brutes a démontré la parfaite redondance géométrique de 1775 bâtiments. Ils ont tous été supprimés. Le mode opératoire a été le suivant :

- Importation de la classe *habitatcub0_0707* dans une base provisoire (*PostgreSQL*) ;
- Réalisation de la requête contenant une fonction spatiale permettant de dresser une liste des identifiants de buildings redondants :

```
select b2.id from public."batiments" b1,
public."batiments" b2
where equals(b1.the_geom, b2.the_geom)
and b1.id<b2.id
```

- Suppression des bâtiments sur base des identifiants listés.

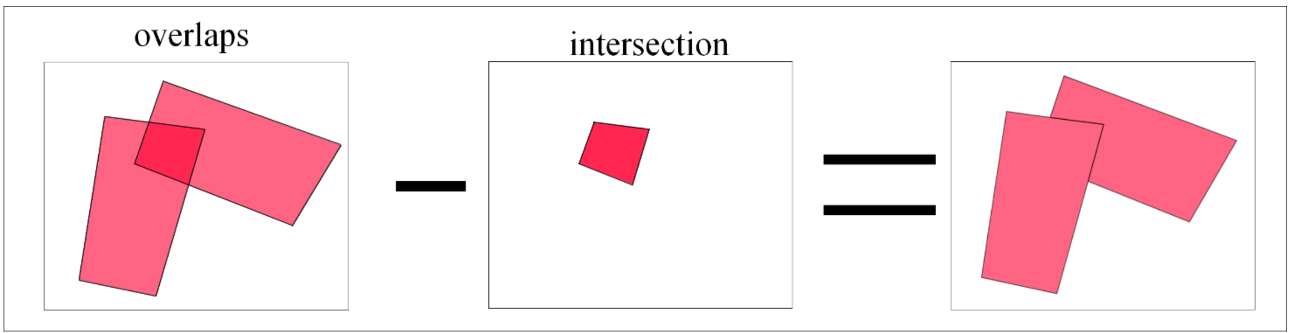
Il est important de spécifier que cette couche de bâtiments ne contient aucune donnée attributaire utile, et que la suppression pure et simple de ces bâtiments n'entraîne aucune perte d'informations.

4.5.3.2 Suppression des intersections

En vue de la suppression des intersections entre bâtiments, la couche a été importée dans une base *PostgreSQL* provisoire, et les étapes suivantes ont été réalisées grâce à des requêtes *Structured Query Language* (SQL):

- Identification des polygones qui s'intersectent (fonction `overlaps(geometry, geometry)`) => *polyoverlaps.shp* ;
- Détermination des géométries des polygones de chevauchement de *polyoverlaps.shp* grâce à `intersection(geometry, geometry) =>` *polyintersection.shp* ;
- Sélection, à ce stade, des identifiants les plus élevés des buildings dans *polyoverlaps.shp* => *polyoverlapsplus.shp* (car les relations inverses sont aussi présentes : si A intersecte B, B intersecte aussi A, d'où présence de deux entités distinctes) ;
- Soustraction des polygones (avec condition d'égalité des identifiants) : *habitatcub0_0707.shp – polyoverlapsplus.shp = BuildingsOK.shp*

Figure 20. Traitements pour suppression des intersections dans la couche Buildings



(Source : DEMONCEAU 2010)

La couche *BuildingsOK* contient cependant des défauts de topologie qui n'existaient pas dans la couche d'origine. La vérification a fait ressortir pas moins de 759 défauts (multi-géométries, polygones auto-intersectants, …). Ce procédé a donc conduit à de nouveaux défauts qui auraient dû être corrigés manuellement étant donné qu'une fonction de nettoyage automatique n'était pas disponible parmi les logiciels utilisés. Devant l'ampleur du travail à réaliser, nous avons décidé de renoncer à cette suppression et nous nous sommes limités à garder les buildings qui s'intersectent en le spécifiant dans le rapport de qualité des données opérationnelles.

4.5.3.3 Complétion de la couche de bâtiments

Une analyse approfondie portant sur les bâtiments au voisinage des point d'affectation des bâtiments a mis en évidence que certains bâtiments présents sur l'image aérienne ne sont pas digitalisés. Les entités manquantes, au nombre de 7 (Cf l'analyse de données), ont été digitalisées manuellement. La couche est rebaptisée *Buildings*.

4.5.3.4 Suppression d'objets non-liés à des bâtiments

Certains objets présents dans différentes classes ne peuvent être reliés aux bâtiments (les objets *cemetery* de *religious institution).* Ils sont donc supprimés.

4.5.3.5 Transfert d'objets entre classes

Certains objets se trouvent dans des classes inappropriées. Par exemple, il est incohérent que « *Pompes funèbres* » se trouve dans la classe « *sécurité publique* ». Le transfert d'entités a été réalisé entre *sécurité publique* et *service public et parapublic*.

4.5.3.6 Fusion des classes d'affectation

Le regroupement de toutes les classes d'affectation en une classe unique facilitera ultérieurement la manipulation dans la base, et allègera le modèle conceptuel. Cette opération n'est réalisable qu'en normalisant préalablement les champs de toutes les classes et en ajoutant un champ qui spécifie le type d'affectation qui est le plus souvent le nom de la classe d'origine (*Religious Institution*, *Public Parapublic Services*, …).

Les champs sont normalisés dans les tables *dbf* de chacun des ensembles de fichiers :

Figure 21. Champs normalisés dans les tables dbf

F_ID,N,16	F_CATEGORY,C,25	F_CODENAME,C,52	F_NAME,C,65	OWNER,C,28

(Source : DEMONCEAU 2010)

Les colonnes superflues sont supprimées et les en-têtes sont renommés suivant la figure 21. Notons que les attributs présentés dans cette figure sont encodés suivant la structure *dbf.*, c'est-à-dire que le nom de l'attribut est suivi de *N* ou *C,* spécifiant le caractère numérique ou textuel des valeurs pouvant être prises, et de leur longueur (prise suffisamment grande pour éviter une troncature).

L'opération de fusion de toutes les classes est ainsi réalisée, et la classe de sortie est nommée *Functions*.

4.5.3.7 Population dans les bâtiments

La couche *population figures* contient un ensemble de points donnant des informations sur la population en place en ces lieux respectivement en 1987, 1992 et 2005. Une couche polygonale aurait été plus appropriée pour représenter les zones de population, mais elle n'est pas disponible, et il n'y a pas, non plus, un point par village ou quartier auquel il serait possible de les y rattacher. Les informations sur la population sont pourtant capitales dans le cas d'un modèle de risque puisqu'il faut savoir « où vit » la population pour estimer le potentiel de pertes humaines en cas d'aléa.

La solution proposée dans ce mémoire est de rattacher ces données de population aux bâtiments « *non affectés* » c'est-à-dire les bâtiments sans fonction et supposés être des habitations. Le but étant de déterminer le nombre d'habitants par maison. Cela peut être réalisé en polygonalisant les points de population (Thiessen), et en divisant ensuite le nombre d'habitants par zone par le nombre de maisons dans cette zone pour avoir une idée de la taille moyenne des ménages.

La toute première étape est donc d'arriver à identifier les habitations parmi les bâtiments présents dans la couche *Buildings* et d'exclure ceux qui font l'objet d'une quelconque affectation. Cependant, à ce stade du mémoire, il n'est pas encore possible de connaître les fonctions des bâtiments. Ces fonctions sont reprises dans une couche de points. Il est possible d'associer une fonction à un bâtiment lorsque le point correspondant tombe dans un bâtiment. Cependant, tous les points ne tombent pas forcément dans un bâtiment et il est difficile dans ce cas de savoir à quels bâtiments ils se rapportent (figure 22).

Figure 22. Indétermination quant aux bâtiments associés à la fonction

(Source : DEMONCEAU 2010)

La partie consacrée à la création des classes d'associations au niveau du modèle logique de données, dans le chapitre 5, apporte une réponse à ce problème. Il sera alors possible d'identifier les bâtiments sans fonction et de poursuivre le traitement envisagé ci-dessus.

Dès lors, pour l'instant, seuls les champs *pop_1987*, *pop_1992* et *pop_2005* peuvent être ajoutés à la table table *dbf* de *Buildings* mais ils ne seront peuplés qu'ultérieurement lorsque toutes les habitations seront identifiées.

4.5.4 Paquetage « Hydrographie »

La classe qui paraît la plus pertinente au point de vue géométrique est *hydro bda rivers*. Elle couvre la même étendue que l'image Ikonos et contient le réseau qui s'y ajuste le mieux. Les autres classes ne contiennent aucune information qui pourrait l'enrichir de sorte qu'aucun traitement n'est à réaliser ici pour optimiser les données retenues.

4.5.5 Paquetage « Relief »

Une première analyse des classes entraîne l'exclusion immédiate de *contour satellite towns* à cause du médiocre intervalle de hauteur entre courbes de niveau (50m). La classe *station escarpement* est également éliminée étant donné que l'identification de l'escarpement scindant la ville en deux parties est facilement reconnaissable dans les autres classes grâce à l'analyse de la densité des courbes de niveau. La précision géométrique des deux classes restantes est exactement la même, et seule *courbes de niveau bamenda* est conservée étant donné sa couverture spatiale égale à celle de l'image aérienne considérée comme fond de référence. La classe d'objets est rebaptisée *Contourlines*.

L'analyse de la qualité géométrique des données discutée dans la section précédente (relative à l'analyse des données) a fait ressortir un décalage d'une soixantaine de mètres entre les couches de courbes de niveau et la photographie aérienne. Un recalage de la couche par transformation affine aurait pu être envisagé pour corriger ces défauts. Cependant, ne connaissant pas la manière dont ces courbes de niveau ont été créées, ni sur base de quelles sources elles l'ont été, il était impossible de déterminer avec précision les paramètres de cette transformation.

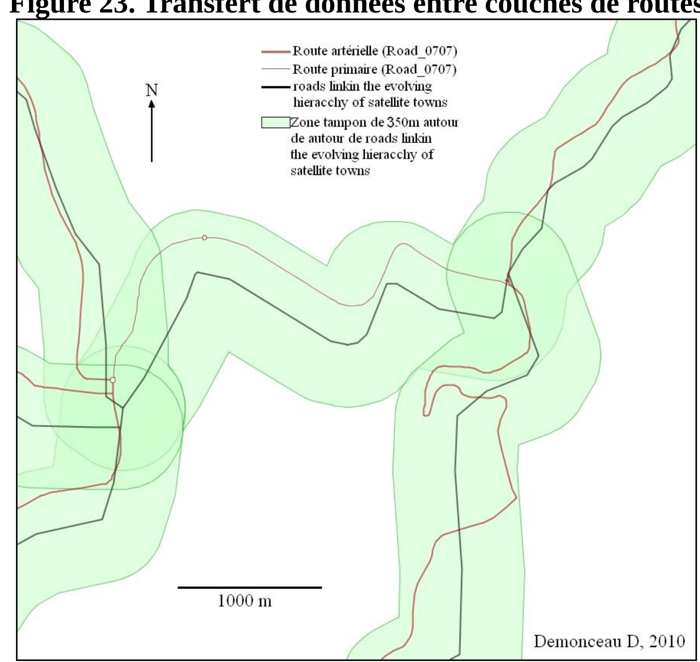
4.5.6 **Paquetage « Voiries »**

4.5.6.1 Routes

Les classes *road network*, *roads linkin the evolving hieracchy of satellite towns* et *roads_0707* contiennent le réseau routier, mais seule cette dernière est conservée en raison de ses meilleures qualités géométriques et de la plus grande densité du réseau couvert ainsi qu'en raison du caractère plus récent du relevé (identifié par le suffixe *0707*).

Cette classe peut être complétée par des informations relatives aux noms de routes contenues dans la classe *roads linkin the evolving hieracchy of satellite towns*. Cette couche contient les routes artérielles et primaires de la ville ainsi que leurs noms (ces noms ne figurent pas dans *roads_0707*). Le transfert de données utiles passe par l'identification des routes homologues dans chacune des couches. Cette tâche n'est pas évidente car leurs tracés diffèrent dans les deux couches (problèmes d'exactitude), comme illustré sur la figure 23.

Figure 23. Transfert de données entre couches de routes



(Source : DEMONCEAU 2010)

Une colonne *ROAD_ID* est tout d'abord créée afin d'engendrer un identifiant non équivoque pour toutes les routes.

Le transfert d'informations est réalisé sur base d'une distance tampon de 350m[9] autour des routes « *primaires* » et des « *artères routières* » de *roads_0707*. Toutes les entités « *route* » de *roads linkin the evolving hieracchy of satellite towns* qui tombent dans cette zone voient leurs informations extraites et introduites dans *roads_0707*. La figure 23 met cependant en évidence la superposition des zones tampon, et il est facile de voir que certaines routes, se trouvant dans ces zones, font l'objet de « *multi-transferts* » d'attributs. Un nettoyage, consistant en une suppression manuelle des données non-voulues, est donc réalisé.

9 Distance déterminée arbitrairement pour être certain que l'identification de toutes les routes homologues soit possible

Les résultats de ces traitements sont réintroduits dans *roads_0707* et nettoyés pour ne garder que les informations utiles. La couche finale est rebaptisée *Roads*. L'ensemble des opérations décrites est repris à la figure 24.

Figure 24. Flux de traitements associés au transfert de données entre les classes de routes

Road_0707
SHP
Roads linkin the evolving hieracchy of settlements
SHP
Transfert d'attributs
(2) dans (1)
(1) < 350m de (2)
Roads2
SHP
Requête attributaire
RT_CLASS = 'ARTILIARY'
Requête attributaire
RT_CLASS = 'PRIMARY DISTRIBUTOR'
Artiliary
SHP
Primary
SHP
Jointure attributaire
Sur base de l'identifiant
Roads
SHP

(Source : DEMONCEAU 2010)

4.5.6.2 Parkings

Deux classes sont relatives aux parkings mais seul *Parking facilities revised* est conservé étant donné son caractère « *révisé* » par rapport à celui de l'autre classe qui semble être plus ancienne.

Les objets de type « *parking* » sont sous forme ponctuelle dans le *shp.* Cela ne représente pas la primitive géométrique adéquate pour décrire les objets. La représentation la plus pertinente serait sous forme polygonale. Une « polygonalisation » a dès lors été réalisée manuellement.

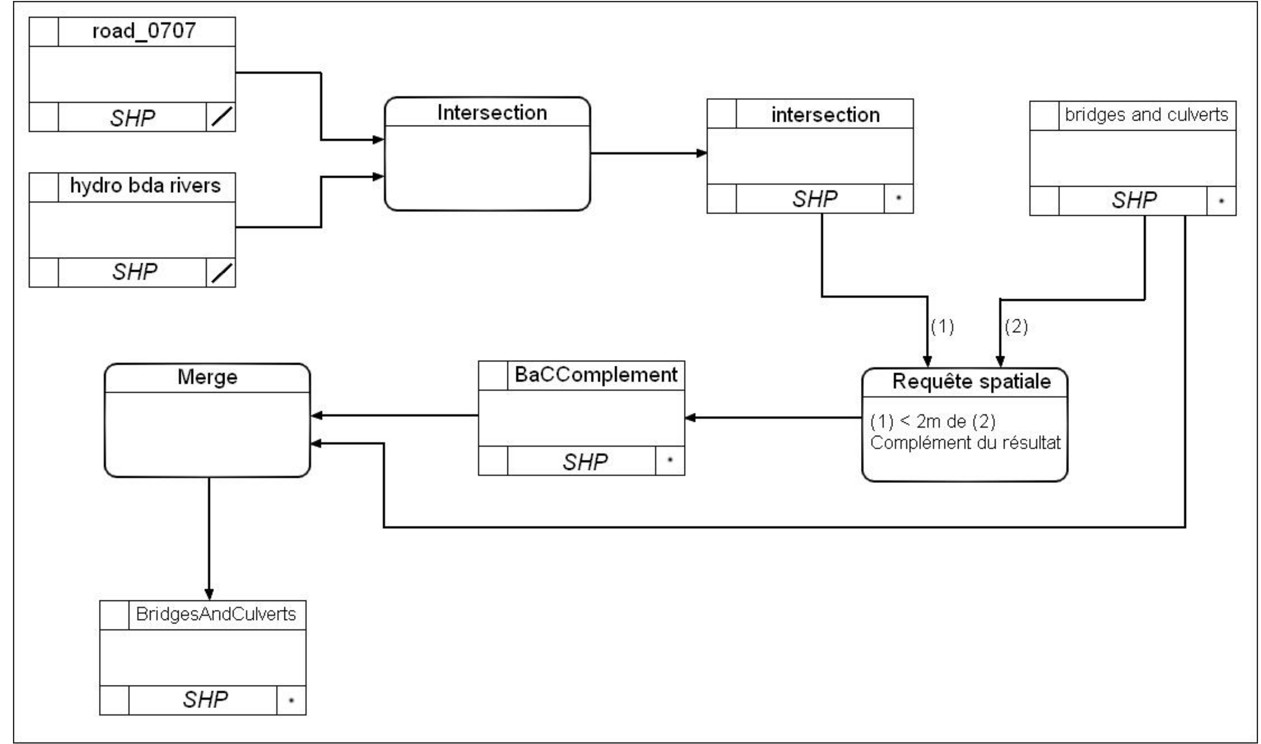
La digitalisation des zones de parkings se base sur l'identification visuelle des emplacements visibles sur l'image aérienne. C'est le seul document disponible pour repérer les limites.

Des polygones sont donc créés et ne contiennent aucune donnée dans la table *dbf* associée. Il faut alors réaliser une intersection avec transfert d'attributs de la couche ponctuelle contenant les informations (type parking, nom et statut) vers la couche nouvellement créée.

4.5.6.3 Ponts et caniveaux

La section relative à l''analyse des données a mis en évidence un manque de 50 points d'intersection entre le réseau routier et le réseau hydrographique. Les traitements réalisés et schématisés à la figure 25 ont permis de créer une nouvelle couche comprenant les ponts complémentaires.

Figure 25. Flux de traitement associés à la création d'une couche contenant les ponts complémentaires



(Source : DEMONCEAU 2010)

Ces « *ponts* » supplémentaires prennent la valeur « *To determine* », pour les raisons citées à la sous-section 4.2.2.3.1. relative au contrôle qualité des données brutes, et sont ajoutés à la couche de ponts d'origine via une fusion. La classe résultante est nommée *BridgesAndCulverts*.

4.6 Renommage

L'analyse de la qualité sémantique a fait ressortir un problème au niveau de la ou des langues utilisées dans les noms de classes et de champs. Le choix s'est finalement porté sur la présentation de tous les noms en anglais, et ce pour deux raisons :

- l'utilisateur principal, demandeur du prototype, NYAMBOD, est anglophone ;
- la ville de Bamenda se trouve dans l'une des deux régions anglophones du Cameroun.

4.7 Catalogue des données

Les données issues des précédents traitements de préparation ont une forme définitive et font partie des données opérationnelles qui intègreront le prototype. Il s'avère dès lors important de créer un catalogue de données afin de (COLLIGNON 2004):

- dresser une liste complète des données utilisées au sein du prototype ;
- offrir une description des données, ainsi que des attributs, des opérations liées et des associations entre elles ;
- déterminer le type de primitive géométrique de chacune des données spatiales ;
- évaluer la qualité sémantique ;
- identifier le créateur du catalogue et le contacter en cas de problème.

Un catalogue fourni avec le produit « *prototype SIG* » renforce le caractère complet du travail et procure une description normalisée des données intégrées. Structuré de manière similaire à la base de données française BDTopo (IGNF 2002) (figure 26), ce catalogue des données opérationnelles se trouve en annexe en format papier. Nous aurions pu implémenter ce catalogue dans un SGBD, cela se fait couramment dans les milieux professionnels. Nous ne l'avons cependant pas fait étant donné que l'exploitation de ce catalogue requiert la connaissance du langage SQL. Pour les non-initiés au langage, tels que NYAMBOD, le format papier est beaucoup plus accessible.

Figure 26. Classification des couches

A	***Voies de communication routière***		
A 1	CriticalAreasAndTroubleSpots	A 2	Roads
B	***Hydrographie***		
B 1	Rivers		
C	***Surfaces d'activités et bâti***		
C 1	BridgesAndCulverts	C 3	Functions
C 2	Buildings	C 4	Parkings
D	***Occupation du sol***		
D 1	Landuse		
E	***Orographie***		
E 1	Contourlines		
F	***Zonages techniques et administratifs***		
F 1	Quarters	F 3	Town
F 2	Subdivisions	F 4	Villages
G	***Objets divers***		
G 1	Bamenda		

(Source : DEMONCEAU 2010)

5 Modélisation des données

5.1 Modèle conceptuel de données

Les données faisant partie de la modélisation sont les 13 *shp* issus de la batterie de traitements d'optimisation, ainsi que l'orthophotographie, document fondamental constituant le fond de digitalisation des données s*hapefile*.

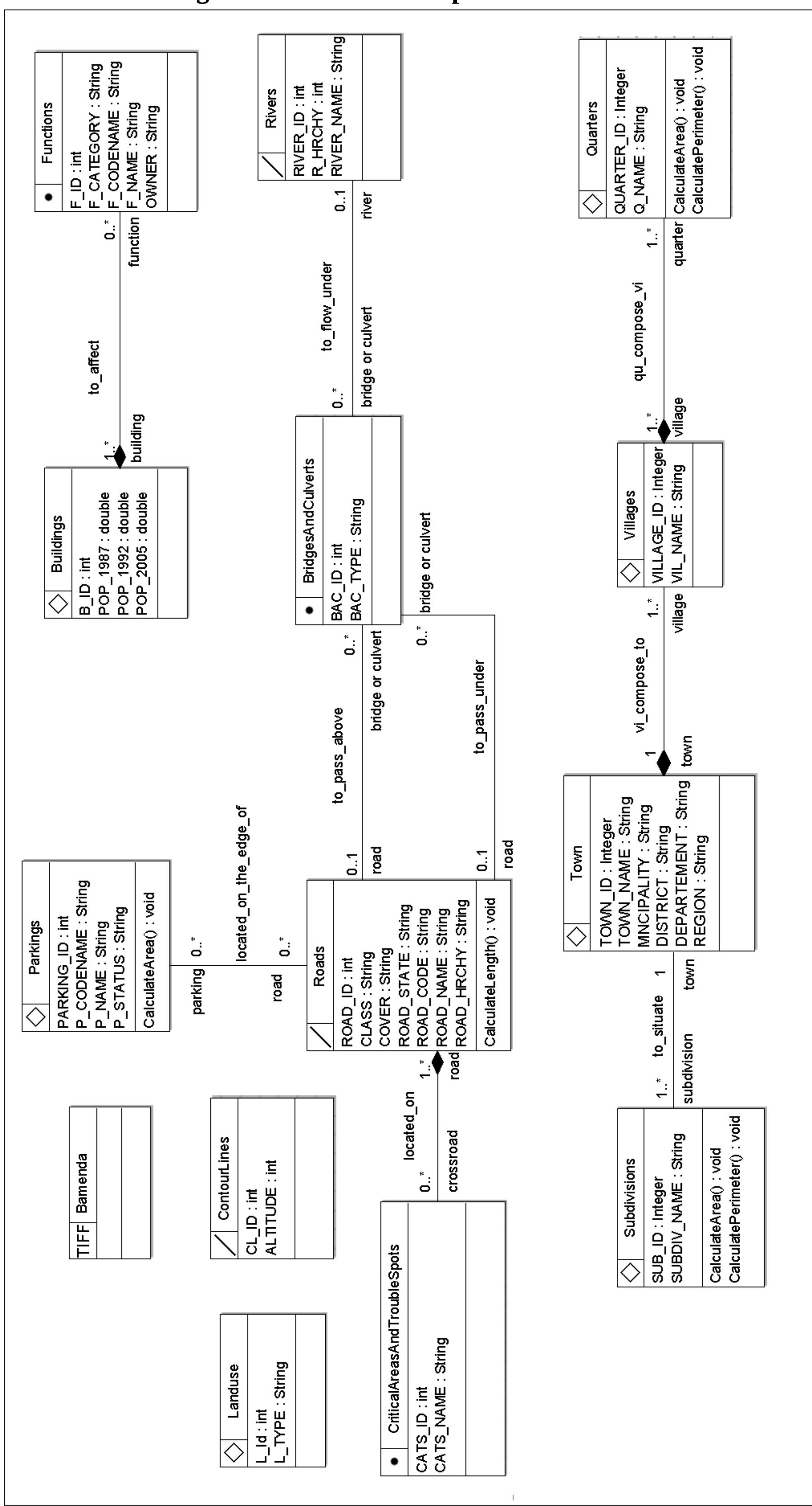
Nous avons dressé, en collaboration avec NYAMBOD, le schéma conceptuel de données qui est repris en figure 27. Il est important de noter que, par suite de leur trop grand nombre, tous les liens spatiaux possibles (intersections, jointures spatiales,…) ne sont pas représentés dans ce schéma conceptuel. Ainsi, bien qu'il soit tout à fait possible de connaître, en réalisant une intersection de couche, dans quel village se trouve tel ou tel bâtiment, ce genre de lien n'est pas représenté ici.

Nous avons également décidé de ne pas représenter les paquetages car, d'une part, ceux qui sont identifiés dans le catalogue contiennent peu de classes, et d'autre part, surcharger le schéma conceptuel rendrait sa lecture moins aisée.

Les classes *Bamenda*, *Landuse* et *ContourLines* sont des classes non liées car il s'agit de classes d'objets « *non-tangibles* » indépendantes les unes des autres.

Les noms des relations inverses ne sont pas mentionnés dans le schéma conceptuel repris à la figure 27 car ils sont implicites. Le choix des types d'association et des types de cardinalité est explicité ci-après.

Figure 27. Schéma conceptuel de données



(Source : DEMONCEAU 2010)

5.1.1 **Associations et cardinalités**

5.1.1.1 located_on

Cette relation est de type *n-n* car une route peut être composée de plusieurs tronçons critiques, et un carrefour est le croisement de plusieurs routes.

Cette association est de type « *composition* » car sans routes, il n'y a pas de tronçons critiques.

5.1.1.2 located_on_the_edge_of

Cette relation est également de type *n-n* étant donné qu'un parking peut être bordé ou non d’une ou plusieurs routes, et une route peut être bordée ou non d’un ou plusieurs parkings.

5.1.1.3 qu_compose_vi

Dans la réalité, un quartier n'appartient qu'à un seul village, et un village peut contenir plusieurs quartiers. Dans le schéma, les cardinalités de cette association ne respectent toutefois pas cette règle « *vraie* » suite à l'inexactitude des polygones de quartiers et de villages. En effet, les bords de ces polygones n’étant pas superposés, certains quartiers peuvent chevaucher plusieurs villages.

La relation est ici également de type « *composition* » car les quartiers « *composent* » les villages, et les villages disparaissent si les quartiers disparaissent.

5.1.1.4 to_affect

Cette relation est de type *n-n* car un bâtiment peut n’abriter aucune fonction (dans ce cas, il s’agit d’une habitation), en abriter une, ou en abriter plusieurs (exemple : banque et assurances dans un même bâtiment). Et une fonction peut être affectée à un ou à plusieurs bâtiments (école, brasserie,...).

Dans cette relation, la cardinalité minimale de *Functions* dans la relation est 1 et non 0 car les fonctions sont toujours affectées à un bâtiment au moins, les fonctions non liées à des bâtiments (exemple : cimetières) ayant été préalablement éliminées de la classe *Functions*.

La relation est de type « composition » car un bâtiment permet l’exercice d’une fonction, et l’absence de bâtiment implique l’absence de fonction.

5.1.1.5 to_flow_under

Une rivière peut ne couler en-dessous d'aucun pont, ou couler sous un ou plusieurs ponts. Et un pont peut ne surmonter aucune rivière, ou en surmonter une. Un pont peut aussi, par exemple, surmonter, une route ou une vallée sèche, d'où la cardinalité minimale de 0, dans l'association pour la rivière.

5.1.1.6 to_pass_above et to_pass_under

Une route peut ne passer au-dessus d'aucun pont, ou passer au-dessus d’un ou plusieurs ponts. Et une route peut ne passer en-dessous d'aucun pont, ou passer en-dessous d’un ou de plusieurs ponts. Le pont peut, quant à lui, ne permettre le passage d'aucune route, ou permettre le passage d'une route, et ne surmonter aucune route, ou en surmonter une. Le pont peut effectivement ne surmonter aucune route, ce qui explique la cardinalité minimale de 0 et, dans ce cas là, c'est une rivière ou une vallée qui est surmontée.

5.1.1.7 to_situate

La ville de Bamenda chevauche trois arrondissements et un arrondissement peut contenir plusieurs villes. Dans notre modèle, un arrondissement ne contient qu'une seule ville, celle de Bamenda, ce qui justifie la cardinalité 1.

5.1.1.8 vi_compose_to

La ville contient un ou plusieurs villages, et les villages ne composent qu’une seule ville. Le terme « village » n'a pas le sens habituellement rencontré. Au Cameroun, villes et villages ne se trouvent pas au même niveau hiérarchique de découpage administratif, et un village est un sous-ensemble de la ville (NYAMBOD 2010).

Cela se traduit également ici par une association de type « *composition* ».

5.1.2 **Vérification du modèle**

Les traitements précédents ont permis de conformer les classes au MCD qui doit maintenant être vérifié. Le manuel d'utilisation du SGBD SYBASE précise pour chaque objet une série de critères sur lesquels doit porter cette vérification. Ils sont repris à la figure 28. Une valeur « *CAT* » indiquée dans la colonne « vérification » signifie que le MCD est déjà conforme par rapport au critère en question car celui-ci est repris dans la norme ISO 19110 qui a été appliquée pour le catalogage des données.

La vérification de la bijectivité de l'association entre deux entités n'est pas applicable ici car notre modèle ne possède pas que des relations de type 1.1. Le paramètre 15 n'est pas non plus applicable car nous avons précisé que les relations inverses sont facilement déductibles à partir de la relation directe. Le paramètre 3 n'a pas été vérifié car sa signification n'a pu être déterminée. Il pourrait s'agir ici d'un paramètre propre au logiciel SYBASE.

Figure 28. Critères de vérification du modèle conceptuel

Objet	n°	Critère	Vérification
Entité	1	Unicité du nom et du code d'entité	CAT
	2	Existence d'un attribut	CAT
	3	Types de données séquentiels	
	4	Existence de relations	CAT
	5	Existence d'un identifiant	CAT
Attributs d'entité	6	Unicité du nom et du code d'attribut d'entité	CAT
Identifiants d'entité	7	Unicité du nom et du code d'identifiant	CAT
	8	Existence d'un attribut	CAT
Associations	9	Unicité du nom et du code d'association	CAT
	10	Association dotée de deux liens au moins	CAT
	11	Association dotée de deux liens avec un lien identifiant	CAT
	12	Association dotée d'un seul lien identifiant	CAT
	13	Association bijective (1.1) entre deux entités	NA
	14	Cardinalité maximale des liens	CAT
	15	Liens identifiants réflexifs	NA
	16	Contrainte d'unicité de nom pour les associations de type n.n	CAT

(Source : d'après SYBASE 2010)

5.2 Modèle logique de données

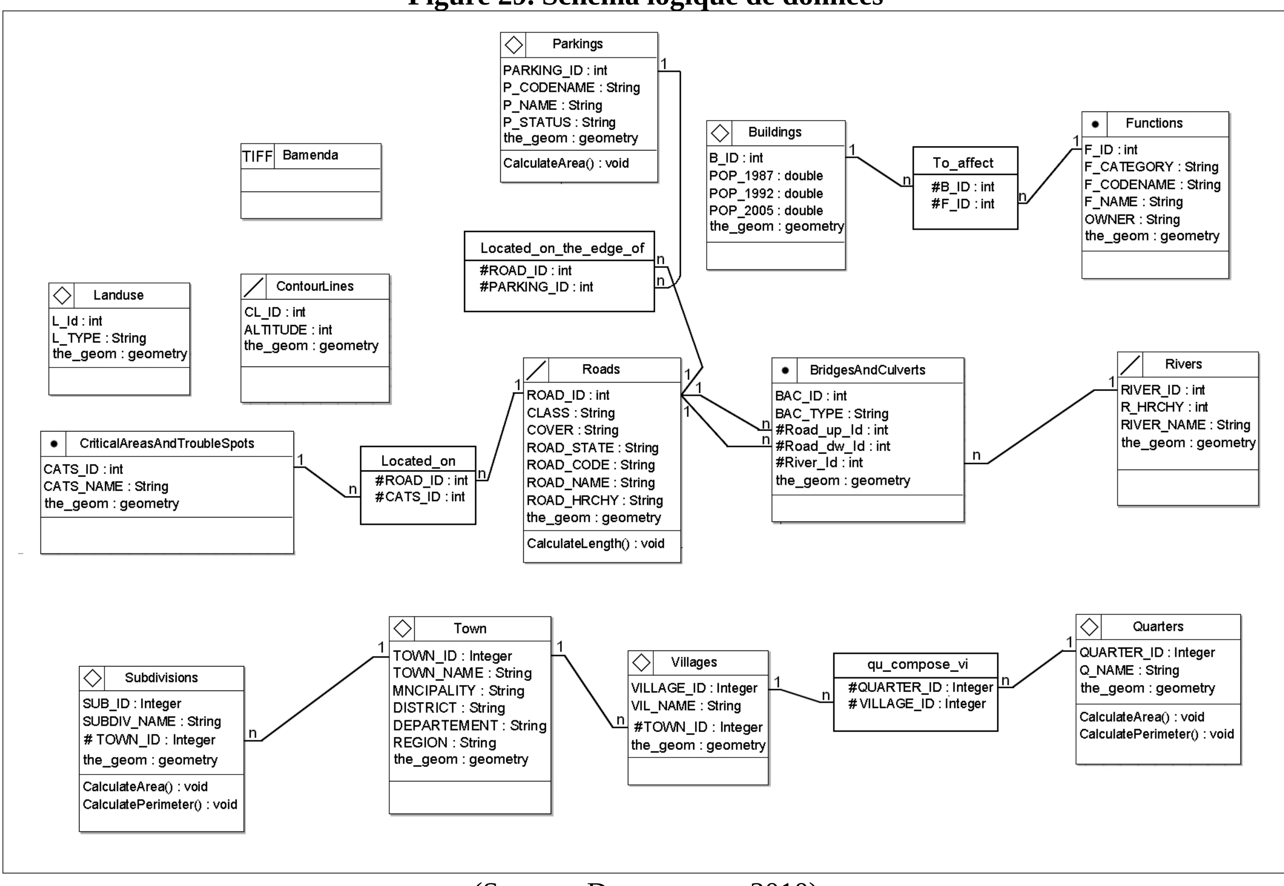
Les liens identifiés au niveau du schéma conceptuel doivent être possibles entre les classes reprises au sein de la base. En d'autres mots, il faut qu'il y ait des champs communs aux classes afin de rendre possibles les jointures non-spatiales entres classes. Une classe participant à une relation de type *1-n*, nécessite l'ajout, sous forme de clé externe, de l'attribut clé de l'autre classe lorsqu'elle se trouve du côté *n*. Une relation de type *n-n* entraîne la création d'une classe d'association composée d'au moins deux champs : les clés des classes participantes.

Les paragraphes qui suivent décrivent les traitements effectués pour créer ces attributs de jointures.

Notons qu'un des traitements concerne la création de la classe d'association qui permet de faire le lien entre bâtiments et fonctions de bâtiments. Rappelons-nous, à cet effet, qu'une section antérieure avait abordé le problème d'intégration de données de population dans la classe *Buildings*. Il nous était alors impossible de réaliser cette tâche car il était nécessaire de connaître les habitations (=bâtiments sans fonction). La dernière section de ce chapitre se consacre à la complétion de la classe *Buildings* car il sera dès lors possible de différencier bâtiments fonctionnels et habitations.

Le schéma logique de données est repris à la figure 29.

Figure 29. Schéma logique de données



(Source : DEMONCEAU 2010)

5.2.1 Création des classes et des relations

5.2.1.1 Association CriticalAreasAndTroubleSports-Roads

Les endroits critiques sur les routes se situent au croisement ou non de deux ou de plusieurs routes. La relation est de type *n-n*, ce qui conduit à la création d'une classe d'association. Elle contiendra deux clés externes faisant référence aux clés primaires des classes *Roads* et *CriticalAreasAndTroublesSpots*.

La meilleure façon de procéder pour déterminer quelles zones critiques font partie de quelles routes est de travailler avec des distances tampons. La topologie médiocre des couches est telle que les points d'endroits critiques ne tombent pas sur des lignes de type route, ce qui exclut une recherche par intersection de couches.

Une distance tampon minimale de 10 mètres autour des points « *zones critiques* » a été retenue afin que toutes les routes concernées y soient incluses. Les attributs sont transférés d'une classe à l'autre et la classe résultante, appelée *Located_on* comme dans le schéma conceptuel, est nettoyée pour ne garder que les champs relatifs aux clés primaires des classes *Roads* et *CriticalAreasAndTroublesSpots*.

5.2.1.2 Association Roads-Parkings

Le traitement relatif à la création de la classe d'association est similaire à celui qui a été réalisé pour la classe précédente. Une distance tampon minimale de 10 mètres autour des bordures de parkings a été retenue pour s'assurer que toutes les routes ayant un accès aux parkings soient prises en compte. La classe résultante *Located_on_the_edge_of* est nettoyée pour ne garder que les clés de *Parkings* et de *Roads* c'est-à-dire PARKING_ID et ROAD_ID.

5.2.1.3 Association Roads_BridgesAndCulverts

La relation entre ces classes est de type *1-n*. La classe *BridgesAndCulverts* accueillera, dans ce cas-ci, un nouveau champ faisant référence à l'identifiant des routes passant au-dessus et en-dessous du pont.

De même, il n'est pas possible de retrouver ces routes par simple intersection de couches étant donné les relations topologiques médiocres existant entre ces couches où les points « *ponts* » ne tombent pas sur les segments « *routes* » (Cf. le chapitre relatif à l'analyse des données). L'association entre ponts et routes a été réalisée via un transfert d'attributs de la classe de routes vers la classe de ponts sur base d'une zone tampon de 5 mètres autour des entités de type ponts.

Le fichier contenant le résultat du transfert a été nettoyé, de manière à ce que seuls les champs *ROAD_UP_ID* et *ROAD_DW_ID* (identifiants respectifs de la route passant par-dessus le pont et de la route passant par-dessous celui-ci) soient ajoutés à *BridgesAndCulverts*.

Il est à noter que la seule manière de déterminer si une route passe en-dessous ou au-dessus d'un pont est d'observer la photographie aérienne. Dans notre cas, aucune route ne passe en-dessous. Le seul intérêt, dans notre travail, de garder un champ *ROAD_DW_ID* est de tenir compte d'éventuelles mises à jour qui pourraient entrainer l'ajout des routes passant en-dessous d'un pont.

5.2.1.4 Association BridgesAndCulverts-Rivers

La même opération que dans le cas précédent a été réalisée pour identifier les rivières coulant en dessous d'un pont (distance tampon de 5m). Un nouveau champ, *RIVER_ID*, est encore ajouté à la classe *BridgesAndCulverts*.

5.2.1.5 Association Quarters-Villages

Le lien entre ces classes est de type « *composition* » car un village est composé de plusieurs quartiers. La relation est de type *n-n*, ce qui se traduit par la création d'une classe d'association via « *transfert d'attributs par intersection de couches* ». La table est nettoyée pour ne garder que *VILLAGE_ID* et *QUARTER_ID*.

5.2.1.6 Association Town-Subdivisions

L'ajout du champ *TOWN_ID* dans *Subdivisions* est réalisé via une intersection entre elle et *Town*.

5.2.1.7 Association Villages-Town

Le type de traitement de cette association est identique au précédent car l'ajout de la clé externe *TOWN_ID* au sein de *Villages* se réalise également via intersection spatiale.

5.2.1.8 Association Buildings-Functions

Le modèle conceptuel spécifie une relation de type *n-n* entre *Buildings* et *Functions*, ce qui signifie qu'un bâtiment peut avoir plusieurs fonctions, et une seule fonction peut affecter plusieurs bâtiments. Cet état de choses se traduit par la création d'une classe d'association *to_affect* contenant les clés *B_ID* et *F_ID*.

Un problème se pose car la primitive géométrique utilisée pour la classe *Functions* est de type ponctuel, et il n'est pas toujours évident de savoir à quel(s) bâtiment(s) un point se rapporte. Beaucoup d'entre eux tombent dans un polygone « *batiment* » mais ce n'est pas le cas pour tous, ce qui exclut la possibilité de procéder à une intersection entre cette couche ponctuelle et la couche *Buildings* pour identifier les bâtiments qui se rapportent aux fonctions. De plus, des mêmes fonctions peuvent affecter plusieurs bâtiments et la simple intersection identifierait au plus un bâtiment.

La solution proposée consiste à définir une distance tampon autour de chacune des entités ponctuelles *Functions* et à attribuer à celle-ci tous les bâtiments « tombant » dans le cercle de rayon égal à cette distance tampon et centré sur le point correspondant. Ce qui implique de se poser la question suivante : *quelles* distances tampons adopter pour *quelles* classes ? Il est clair que la distance varie en fonction du type d'affectation car, par exemple, une entité « *industrie du bois* » rassemble probablement plus de bâtiments qu'une entité de type « *pharmacie* ».

Une première classification a lieu afin de rassembler tous les types d'entités ponctuelles présentant potentiellement les mêmes distances tampons. Cette classification est arbitraire car elle repose uniquement sur la représentation que nous, concepteurs, avons de ces entités. Et il est probable qu'un autre concepteur adopterait une autre classification.

La classification adoptée est celle qui est basée sur les classes élémentaires qui ont donné naissance, par fusion, à la classe *Functions*. Certaines d'entre elles ont cependant encore été subdivisées en sous-classes afin de tenir compte du caractère hétérogène des objets qu'elles contiennent. Un exemple parmi d'autres est la classe *Health* qui regroupe tant des objets « *hôpital* » que des objets « *pharmacie* ». Il est clair que la probabilité qu'un hôpital soit composé de plusieurs bâtiments est plus élevée que dans le cas d'une pharmacie. Les classes retenues, qui sont au nombre de 14, sont les suivantes :

Figure 30. Classification des classes sur base de distances tampons à priori similaires

Type	*Sous-type*	*Distance tampon*
Education		?
PublicParapublicServices		?
PublicSecurity	Camp militaire	?
	Autres	?
ReligiousInstitution		?
Health	Pharmacy	?
	Others	
Telecommunication		?
EquipmentHotelAndTourists	Snacks bars/Inns/Restaurant	?
	Hotel/Chieftainship	?
EquipmentTrader		?
Finance		?
Industry	Brewery/Chemics/FoodTransfo/IndustrialWarehouse/MecaMeta/WoodTransfo	?
	Others/PrintingOff/PublicWorks/Textil	?

(Source : DEMONCEAU 2010)

Il reste maintenant à déterminer une valeur chiffrée en mètres pour les différentes distances tampons. Une distance tampon autour d'un point représente, en fait, le rayon d'un cercle tracé autour de celui-ci.

La méthodologie poursuivie consiste à estimer les distances entre les objets ponctuels de *Functions* et les bâtiments de *Buildings* pouvant potentiellement faire partie de ces affectations. Nous nous aidons également de l'image aérienne pour une meilleure identification des bâtiments. Prendre dans chacune des classes un nombre aléatoire de points et estimer grossièrement les distances tampons, est cependant purement subjectif et non reproductible, ce qui est non-souhaité.

La solution retenue est de déterminer un rayon moyen sur base d'un échantillon d'objets par classe. Selon que la population de la classe est inférieure ou supérieure à 100, respectivement 10 ou 15 points sont tirés au sort selon un échantillonnage aléatoire simple. Les bâtiments pouvant potentiellement faire partie de l'affectation du point sont identifiés visuellement et enveloppés dans un polygone convexe. Le rayon retenu est celui du cercle ayant la même aire que ce polygone convexe. Cette tâche est réalisée pour chacun des individus de l'échantillon, et les rayons sont moyennés pour chacune des classes.

Ces rayons moyens ne peuvent cependant pas être retenus comme distances tampons pour l'ensemble des classes dont sont issus les échantillons car trop nombreux seraient les bâtiments non pris en compte. C'est surtout vrai pour les entités d'affectation caractérisées par des polygones convexes de grandes dimensions.

Le fait d'augmenter les valeurs des rayons retenues augmente le risque d'inclure des bâtiments non-affectés, ce qui est surtout vrai pour les affectations caractérisées par de petits polygones convexes. Par contre, il réduit le nombre de bâtiments ignorés pour les affectations caractérisées par de grands polygones convexes. Il faut donc trouver un juste milieu. Rappelons que, dans le cas de la gestion des risques naturels et de l'identification des zones où vivent les populations, il vaut mieux intégrer trop de bâtiments dans l'analyse plutôt que trop peu. Il n'est donc pas préjudiciable d'augmenter la valeur du rayon moyen.

L'idée est de déterminer un intervalle de confiance autour de la moyenne pour chacun des échantillons et ne garder que la borne supérieure. L'intervalle de confiance considéré comme acceptable a été fixé à 95% autour de la moyenne de chacun des échantillons. Le nombre d'objets dans chacune des classes est relativement petit et ne varie que de 25 pour *PublicSecurity* à 178 pour *PublicParapublicServices*, ce qui pousse inévitablement à utiliser une variable *t de student* comme coefficient multiplicateur de l'écart-type de l'échantillon de rayons.

La borne supérieure de l'intervalle de confiance à 95% autour de la moyenne vaut :

$$\sup(R) = R + t * \frac{\sigma}{\sqrt{n}}$$

où σ est l'écart-type de l'échantillon, où *t* est la variable de *student* à 9 ou 14 degrés de liberté et vaut respectivement 2,262 et 2,145 et où n est la taille de l'échantillon.

Ces calculs ont été réalisés pour toutes les classes, et les résultats sont illustrés dans à la figure 31.

Figure 31. Classes d'affectation et distances tampons (bornes supérieures des intervalles de confiance des échantillons)

Classe	***R (m)***
Educations	53,972
PublicParapublicServices	27,319
PublicSecurity	29,327
ReligiousInstitution	27,989
Health	31,003
Telecommunication	22,737
Snacks bars/ Auberges/Restaurants	13,762
Hotel/Chieftainship	33,498
EquipmentTraders	31,870
Finances	17,758
Brewery/Chemics/FoodTransfo/IndustrialWarehouse /MecaMeta/WoodTransfo	55,986
Others/PrintingOff/PublicWorks/Textil	23,445
Pharmacy	15,571
Military camp	51,352

(Source : DEMONCEAU 2010)

Il est intéressant d'analyser ce tableau afin d'évaluer le degré de pertinence des résultats obtenus. Il en ressort que certaines classes ont des distances tampons similaires comme, par exemple, les classes *PublicParapublicServices* et *ReligiousInstitution*, ou encore les classes *Telecommunication* et *Others/PrintingOff/PublicWorks/Textil*. Ces classes auraient pu être regroupées en une seule pour, notamment, alléger les traitements liés à la recherche des distances tampons.

Cependant, ces possibilités de regroupement de classes sont multiples, et toutes n'ont pas été investiguées. Et la présence de distances très similaires n'est pas, en soi, préjudiciable. La seule raison de procéder à un meilleur regroupement aurait été d'alléger les traitements liés à l'échantillonnage et de gagner du temps à cette étape.

5.3 Complétion de la classe Buildings

Lors de la construction des classes, dans le chapitre 4 relatif à l'analyse et à la préparation des données, il n'a pas été possible de déterminer le nombre d'habitants par maison faute d'identification des habitations dans la classe *Buildings*. Cette tâche est, dès à présent, réalisable car il est maintenant possible de déterminer les bâtiments ayant une fonction et, par conséquent, ceux n'en ayant pas qui seront considérés comme habitations.

Il faut donc tout d'abord identifier les habitations en excluant de la table *dbf* de *Buildings* toutes les entrées référencées par la classe d'association *To_affect*. La classe contenant ces bâtiments non affectés est appelée *BuildingsNoFunctions*.

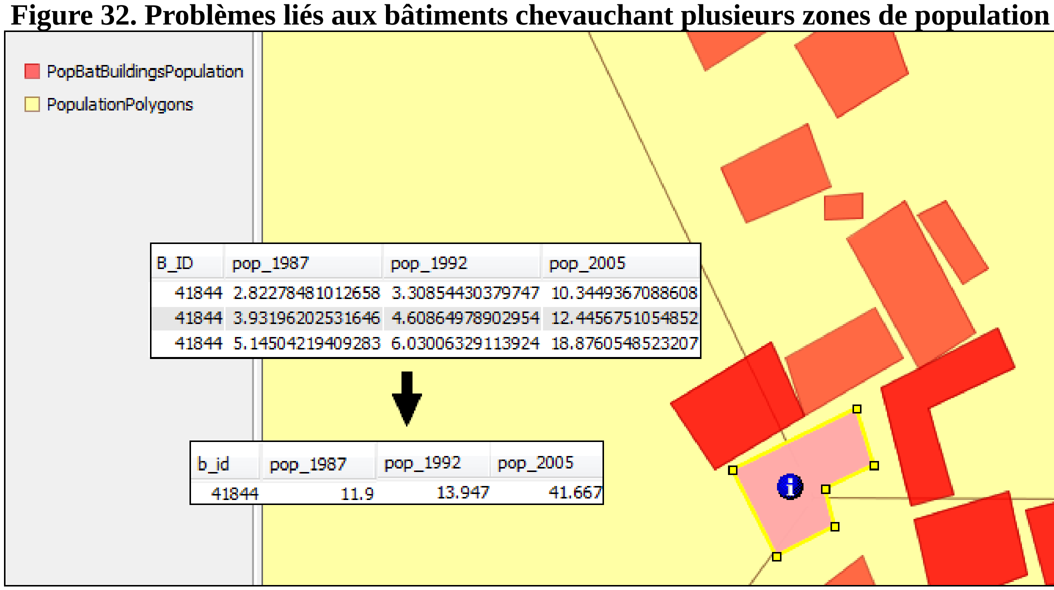
La couche *population figures* fournit, comme cela a déjà été dit, des points informant sur les populations vivant en ces points pour les années 1987, 1992 et 2005. La représentation ponctuelle n'est cependant pas adéquate à l'exploitation. L'idée est de transformer ces entités ponctuelles en entités zonales en procédant à une tessellation de Thiessen. Le résultat obtenu est stocké dans la couche *PopulationPolygons*.

Les couches *BuildingsNoFunctions* et *PopulationPolygons* sont ensuite intersectées avec transfert des attributs de la deuxième couche vers la première. Le nombre de bâtiments par « zone de population » est compté et le nombre d'habitants par maison est calculé via le rapport entre la population d'une zone et le nombre de maisons.

$$\text{taille moyenne des ménages} = \frac{\text{habitants par zone}}{\text{nombre de bâtiments par zone}}$$

Cette intersection avec transfert d'attributs entraîne cependant des redondances de bâtiments en ce qui concerne ceux qui se trouvent à cheval sur deux ou plusieurs zones. Ces bâtiments présentent donc une entrée dans la classe pour chaque zone dans laquelle ils se trouvent (figure 32). Il faut donc remédier à ce défaut en procédant à une « union par valeur d'attribut » avec addition des valeurs de champs. Cela signifie concrètement que les doublons portant sur l'identifiant du bâtiment, *B_ID*, seront fusionnés, et que les autres champs, *POP_1987*, *POP_1992* et *POP_2005*, seront additionnés. Le fichier *BuildingsNoFunctionPop* contient donc, suivant nos hypothèses, toutes les habitations, sans doublons, et les indications sur le nombre de personnes y demeurant respectivement en 1987, 1992 et 2005.

Figure 32. Problèmes liés aux bâtiments chevauchant plusieurs zones de population



(Source : DEMONCEAU 2010)

Les valeurs de tailles de ménage par habitation sont à nuancer pour plusieurs raisons :

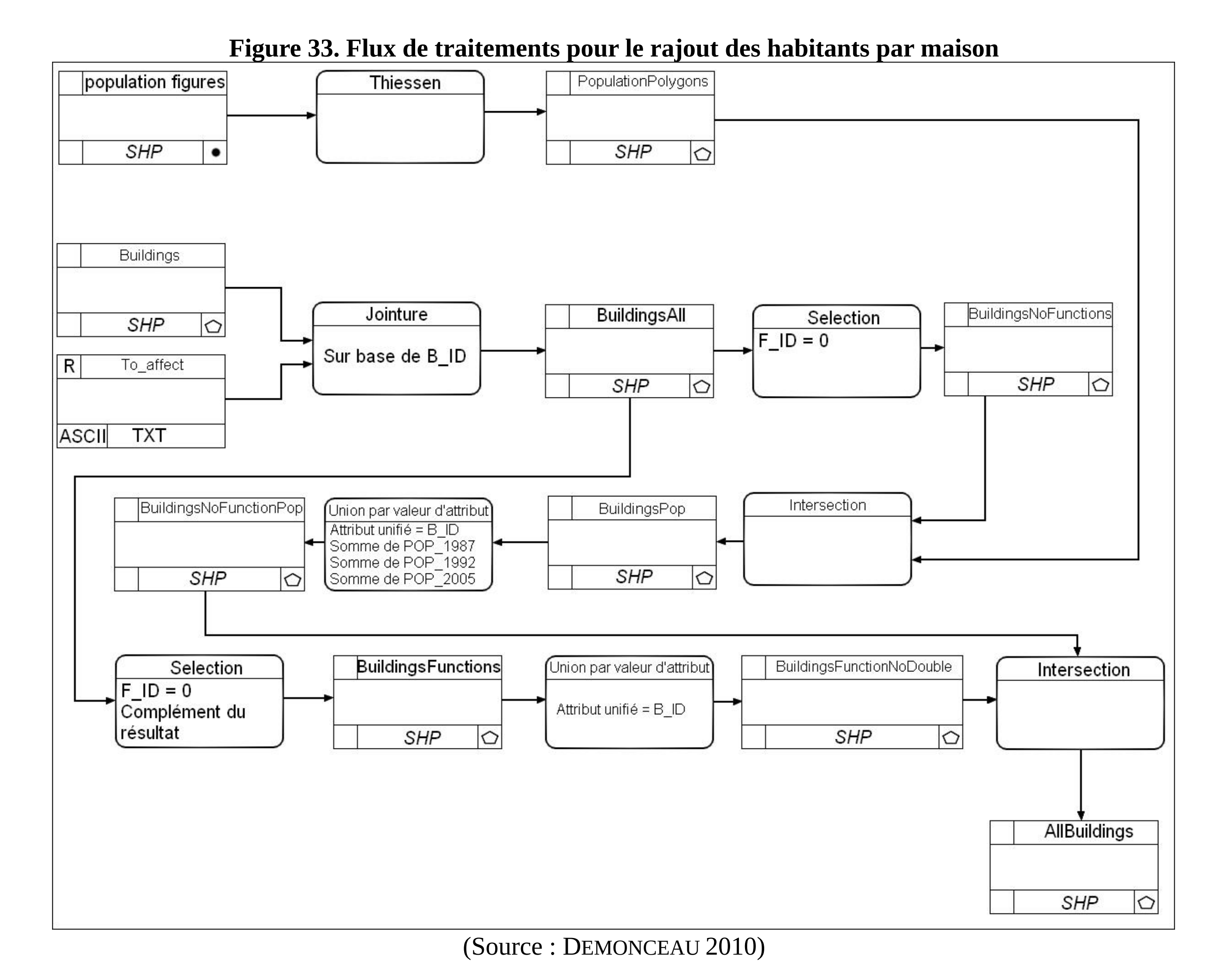
- les données de population sont issues d'un recensement de 1976 et les valeurs pour les années 1987, 1992 et 2005 sont obtenues grâce à une estimation basée sur les taux d'accroissement de population cités par NYAMBOD, la source de ce taux d'accroissement n'a cependant pas pu être déterminée ;
- les tailles des ménages sont calculées en tenant compte d'un nombre identique de bâtiments (+/- 43 000 en 2006) pour les différentes années, ce qui n'est pas réaliste.

Les bâtiments sans fonction doivent être réintégrés au sein de la classe contenant tous les bâtiments qui ont une fonction. Il existe plusieurs manières de le faire, mais la technique retenue consiste à fusionner cette classe avec les bâtiments ayant une entrée dans la classe *To_affect*, et qui ont donc une fonction.

La classe finale *AllBuildings* contient donc tous les bâtiments avec, pour les bâtiments n'ayant pas de fonction(s), indication sur le nombre de personnes y vivant.

L'ensemble des opérations réalisées est résumé à la figure 33.

Figure 33. Flux de traitements pour le rajout des habitants par maison



(Source : DEMONCEAU 2010)

6 Prototype avec data loading

6.1 Modèle physique

Selon CALKINS (1997), la réalisation proprement dite de la base de données s'effectue, entre autres, par la traduction du modèle logique en modèle physique Cette étape est la traduction du modèle logique en modèle physique. Il faut toutefois préalablement créer la base de données.

D'un point de vue pratique, il est également très important d'évaluer les capacités des logiciels et du matériel informatique afin qu'ils puissent répondre aux besoins formulés. Cette étape est d'autant plus nécessaire dans le cas de conception SIG en pays en voie de développement car le matériel informatique y est en général plus précaire.

6.1.1 Évaluation du matériel informatique

CALKINS (1997) suggère de vérifier les aspects suivants pour évaluer les capacités des *hardware* et *software* :

- les fonctionnalités du software ;
- les performances dépendant de :
  1. la manière dont est conçu le logiciel ;
  2. la vitesse de l'ordinateur ;
- l'expansibilité du logiciel et du matériel informatique ;
- la gestion des licences, qui dans le présent cas, n'est pas à prendre en compte étant donné que seuls des produits *OpenSource* sont utilisés ;
- les capacités du système d'exploitation ;
- les processeurs du hardware ;
- l'espace-disque sur le disque dur ;
- l'aspect « *communication* » qui permet l'accès à la base à partir de postes externes.

6.1.1.1 Hardware

Le prototype sera implémenté sur un poste serveur de l'Université de Liège localisé au bâtiment B5 et, plus précisément, à l'étage de l'Unité de Géomatique. Il est en fonctionnement 24 heures sur 24 et 7 jours sur 7 de manière à ce que les utilisateurs autorisés puissent y accéder en permanence.

Ce poste héberge déjà de nombreuses bases de données mais le disque dur possède encore un énorme espace libre. Il ne devrait donc pas y avoir de problème pour ajouter une nouvelle base au SGBD existant.

6.1.1.2 Software

Le SGBD *PostgreSQL*, enrichi avec *PostGIS*, installé sur le poste serveur héberge bon nombre de serveurs exploités par le personnel universitaire. Cet environnement est performant et donc conforme à nos besoins.

6.1.2 **Création de la base de données spatiale**

La base de données est créée à partir des données disponibles et identifiées au niveau conceptuel. Un nouveau nom d’utilisateur est créé à distance dans *PostGIS* et un mot de passe doit lui être attribué. Cet utilisateur sera le gestionnaire du prototype SIG. La connexion à distance requiert la connaissance de l'adresse *ip* du poste serveur. Cette adresse est unique pour chaque poste serveur et permet d'identifier de manière non équivoque un poste connecté sur la toile du web. Le poste serveur du B5 qui sert de support à l'implémentation a l’adresse suivante :

`139.165.44.38`

La base de données est ensuite créée « *physiquement* » via l'interface graphique *PgAdmin III*. Les paramètres du compte qui ont été définis de manière explicites sont les suivants :

Nom d'utilisateur :	`gisbamendauser`
Mot de passe :	`riskmodelbamendaGIS`
Nom de la base de données :	`gisbamenda`

Notons qu'il faut rajouter la ligne suivante dans le fichier de configuration *pg_hba.conf* de *PostgreSQL* afin de permettre la connexion à distance de cet utilisateur :

```
host  gisbamenda    gisbamendauser     139.165.44.38  0.0.0.0 md5
```

'host' est le type de connexion, suivent ensuite le nom de la base et celui de l'identifiant autorisé. L'adresse ip, son masque ainsi que le type de cryptage du mot de passe complètent la ligne.

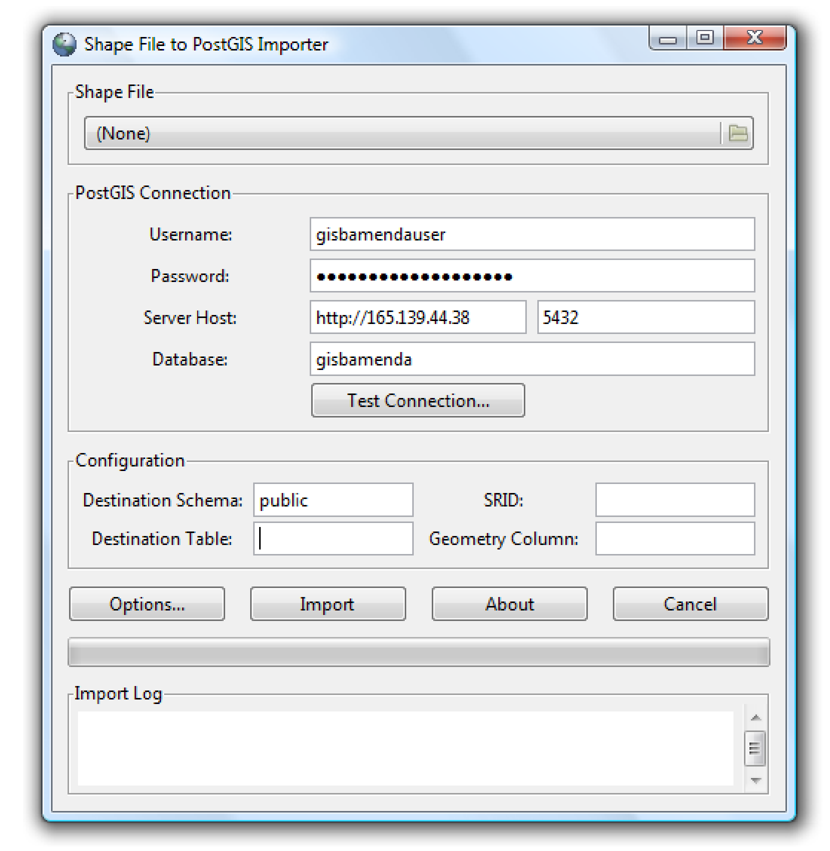
6.1.3 **Implémentation dans le SGBD**

L'implémentation au niveau physique consiste à créer les tables spatiales et non-spatiales et à les peupler avec les données. La création ainsi que le peuplement se font en une seule et même étape grâce au *plugin* « *PostGIS shapefile and DBF loader* » de *PostGIS* pour les données *shp* et *dbf*. L'image aérienne fait l'objet d'un traitement spécial lors de l'implémentation. Le *plugin* est relativement récent (février 2010) et allège considérablement le travail. L'introduction, parfois compliquée, de requêtes SQL n'est donc plus indispensable. L’inconvénient de ce *plugin* est que les phases de création des tables et de peuplement de celles-ci sont réunies en une seule étape. Il est cependant préférable de les réaliser en deux étapes distinctes et d'intercaler entre celles-ci une étape de définition des contraintes d'intégrité de la base pour tester la cohérence entre les données et la structure de la base.

D'après WIKIPEDIA (2010), le *dataloading* comprend les opérations d'extraction, de transformation et de chargement des données. Les deux premières opérations ont été réalisées dans les chapitres précédents ; seul le chargement des données reste à faire. Cette étape constitue la dernière phase d'implémentation au niveau physique de la base de données. Une conversion de données dans un autre format n'est pas nécessaire étant donné que les *shp* et l'image *tif* sont implémentables dans le SGBD *PostGIS*.

Le lancement du *plugin* ouvre tout d'abord une fenêtre (figure 34) dans laquelle l'utilisateur doit s'authentifier et spécifier l'adresse du serveur ainsi que son port d'écoute (qui est 5432 par défaut). Un test de connexion peut être réalisé et est même fortement conseillé pour éviter de perdre inutilement du temps à remplir les autres champs et à chercher le fichier dans le bon dossier (ce qui peut être fastidieux dans le cas d'un poste de travail mal ordonné). Il faut ensuite spécifier l'emplacement de la table dans la base de données (=*Destination Table*) et le paquetage (=*Destination Schema*). La suite des opérations diffère selon que la donnée chargée est spatiale ou attributaire.

Figure 34. Interface de chargement de données



(Source : d'après POSTGIS 2010)

6.1.3.1 Tables spatiales

PostGIS tient compte de la spatialité des données en ajoutant un champ (dont le nom doit être donné (=*Geometry Column*)) contenant la description de la géométrie des entités en format *Well-Known Binary* (WKB) dans la table.

Les données spatiales sont référencées dans un plan de projection. Le champ texte SRID (*Spatial*

Reference Identifier) de la fenêtre permet de spécifier un identifiant de ce plan de projection qui, dans notre cas, vaut 32632 (=UTM ZONE 32 N à partir du datum WGS84). Cet identifiant, qui renvoie directement aux paramètres associés, est très important pour l'exportation des données et leur utilisation à partir d'un poste client hébergeant une ou des applications car il faut que ces applications puissent positionner correctement les données spatiales dans l'espace.

Par défaut, un index spatial de type *R-tree-over-GiST*[10] est créé pour chacune des entités. Cela permet de réaliser une recherche plus rapide sur les données géométriques lors des requêtes.

Lorsque tous les paramètres sont correctement encodés, il faut rechercher le fichier *shp* à charger sur le disque dur. L'étape est validée en appuyant sur le bouton « *import* ». Cette action déclenche les deux étapes suivantes :

– le processus de création de la table au sein de la base. Les entêtes des fichiers *dbf* associés aux *shp* sont extraits et permettent d'attribuer des noms aux champs et de connaître le type de données contenues dans ces champs.

```
CREATE TABLE NomDeLaTable(

   gid           integer,
   attribut01    typeDeVariable,
   attribut02    typeDeVariable,
   ...
   the_geom      geometry
)
```

  Un champ `gid` est automatiquement créé à chaque importation. Il s'agit de l'identifiant unique attribué à chaque tuple par *PostGIS*. La définition du champ contenant la géométrie des objets complète le processus de création de la table.

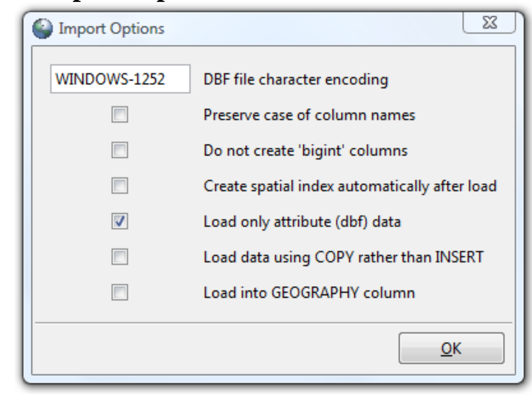
– le peuplement de ces tables. Les informations attributaires sont extraites une à une de la table dbf et insérées dans les champs adéquats grâce à la commande `insert`. Le remplissage de la colonne `the_geom` est réalisé en extrayant les géométries du *main file shp*. Elles sont importées en format WKB (l'équivalent « *texte* » est le WKT). Ce format OGC est utilisé pour le transfert et le stockage des données géométriques sous forme de flux binaire *Binary Long Object* (BLOB). La structure d'un WKB est composée d'un octet d'ordre suivi de 4 octets spécifiant le type de WKB et de la liste des coordonnées X et Y des objets codées chacune sur 8 octets. La correspondance entre une ligne d'information du *dbf* et sa géométrie associée est réalisée à l'aide du fichier d'index spatial *shx*.

10 Plan d'indexation en vigueur dans *PostGIS* depuis avril 2004

6.1.3.2 Tables attributaires

Les documents contenant les informations sur les classes d'association sont issus de traitements opérés sur les données géographiques de format *shp*. Ces documents sont donc aussi en format *shp*. Les géométries contenues dans ces classes sont redondantes car elles sont également contenues dans d’autres classes. Dès lors, le chargement du *shp* complet intégrerait des informations inutiles, or seules les données attributaires doivent être chargées.
Les informations utiles se trouvent dans le fichier *dbf* associé au *main file shp*. Dans le module de chargement, l’une des options reprises en figure 35 permet de charger uniquement les tables *dbf*.

Figure 35. Options possibles lors de l'utilisation du plugin



(Source : d'après POSTGIS 2010)

Il est inutile de spécifier un nom pour la colonne géométrique de la table à créer et il ne faut pas spécifier un SRID car il s'agit de données non spatiales.

La manière de procéder lors de l'importation du *dbf* est sensiblement similaire à l'importation de données spatiales à la seule différence près qu'il n'y a aucune colonne géométrique à créer ni à peupler.

6.1.3.3 Table image

L'intégration de la photographie aérienne au sein de la base est réalisée en deux étapes contrairement à l'intégration des formats *shp*. La création de la table et le chargement des données sont ici deux étapes clairement distinctes. Ce qui constitue un bel exemple didactique de la façon de réaliser de telles tâches sans l'aide du *plugin*.

La photographie aérienne n'est pas en format *shp* mais bien en format *tif*. La méthodologie adoptée consiste à l'intégrer dans la base sous forme d'un objet large (BLOB) via la définition d’une variable *Object IDentifier* (OID). Les octets qui composent l'image sont convertis en flux de données qui sont introduits dans cette variable.

La table qui contiendra cette image a été créée via la commande suivante insérée dans l'éditeur de requêtes :

```
CREATE TABLE NomDeLaTable(
  "name" character varying(20) NOT NULL,
  image oid NOT NULL,
)
```

Les données contenues dans *bamenda.tif* sont transformées en flux binaire et insérées dans la colonne

'*image*' de la table *ortho_image* via la commande :

```
insert into image values ('Bamenda', lo_import('D:/.../bamenda.tif'))
```

La réponse de cette requête est un nombre entier qui correspond à l'identifiant de l'objet donné par *PostGIS*. Cet identifiant permet d'accéder ultérieurement à l'objet (POSTGRESQL 2010).

6.1.3.4 Définition des contraintes d'intégrité

Il reste une dernière étape à réaliser afin d'assurer l'intégrité de la base de données lors de changements de type « *mise à jour* », « *suppression* » ou encore « *création* » au niveau des attributs, des champs ou des tables. Cette étape consiste à « *fixer* » les liens qui existent entre les données et qui ont été identifiés au niveau du schéma conceptuel. Il faut que ces liens persistent et ne se perdent pas en cas de changements. Cette condition est réalisée grâce à l'introduction des « *contraintes d'intégrité* » qui sont des conditions à respecter à tout moment par les données (WOLPER 2007).

Il existe plusieurs sortes de contraintes d'intégrité (liste non-exhaustive) :
- les contraintes d'intégrité locales qui consistent en la définition de clés primaires ;
- les contraintes d'intégrité de référence qui consistent en la définition de clés étrangères ;

Les premières se rapportent uniquement aux tables spatiales tandis que les deuxièmes portent sur les tables liées à d'autres.

6.1.3.4.1 Contraintes locales

Par définition, une clé primaire est un champ ou un ensemble de champs qui permettent de faire référence de manière non équivoque à un enregistrement dans une table. De toute manière, lorsqu'il y a répétition d'une valeur pour les clés, un message d'erreur est retourné lors de la définition de la clé primaire. Toutes les classes ont, par construction (la section 4.5.), un tel champ et ils portent tous le préfixe « Id ».

La spécification, au niveau de la base, d'une clé primaire doit être explicite pour chacune des tables et est réalisée sous *PgAdmin III* en se rendant dans le menu « *contraintes* » de la table correspondante.

6.1.3.4.2 Contraintes d'intégrité de référence

Un attribut d'une table est une clé étrangère lorsqu'il fait référence à un champ d'une autre table. Le champ référencé doit obligatoirement être une clé primaire ou respecter la contrainte d'unicité. La définition d'une clé étrangère permet de s'assurer que la valeur du champ référencé existe bel et bien dans la table référencée. En l'absence de clé étrangère, une perte d'intégrité pourrait tout de même être amenée par l'ajout de valeurs n'existant pas dans le champ référencé et une incohérence ou une erreur pourrait apparaître lors d'une tentative de jointure sur l'une de ces valeurs.

6.2 Validation des données opérationnelles

Les données ont subi un grand nombre de traitements afin de supprimer les erreurs grossières qui existaient dans les géométries et les attributs, et afin de les conformer à la norme ISO 19110. Nous n'avons cependant pu corriger tous les défauts et ceux qui persistent doivent être identifiés dans le rapport de qualité des données opérationnelles.

La validation des données opérationnelles est réalisée à l'aide de requêtes SQL, intégrant des fonctions spatiales ou non, opérées sur les données à partir de l'interface de *PgAdmin III*.

La validation porte sur plusieurs points : le comptage du nombre d'entités par couches, le nombre de points des géométries des objets, le nombre d'intersections, le nombre de multi-géométries par couche, le nombre d'auto-intersections, les redondances géométriques et attributaires, les parfaites redondances ; ainsi que sur la vérification de la fermeture des polygones. En cas de besoin, la validation peut être accompagnée d’un commentaire.

Les deux premiers points sur lesquels porte cette validation donnent une première estimation de la quantité d'informations contenues dans chacune des couches et informent sur l'éventail de variabilité des formes géométriques. La recherche des redondances attributaires est réalisée sur tous les attributs, hormis les attributs de clés, vu que celles-ci sont uniques. Cette recherche n'est pas réalisée pour les classes où l'unicité des informations n'est pas obligatoire telles que la classe *Landuse*, par exemple, où plusieurs polygones peuvent être de même affectation. Le test de recherche de « *parfaites redondances* » est positif lorsque deux entités différentes sont identiques aux niveaux géométrique et attributaire.

Nous avons constaté que la couche *Buildings* est la plus chargée avec 43.236 bâtiments différents. Aucune redondance géométrique n'y a été détectée mais l'analyse nous a indiqué qu'il y avait 1.216 intersections. Pour rappel, une tentative de correction automatisée avait échoué lors de l'étape de préparation des données (sous-section 4.5.3.2).

Au niveau de la couche *Buildings*, il est intéressant de noter que les bâtiments sont composés d’un nombre de points variant de 4 à 21. Les trois entités « bâtiments » comportant 361, 102 et 73 points étaient simplement des ronds-points que nous avons décidé d’écarter de la couche de données.

La couche *Functions* est caractérisée par une forte proportion de redondances au niveau attributaire. Quelque 97 points sont en effet exactement les mêmes à l'identifiant près. Cela fait approximativement 10 % de redondance. Il faut savoir que plusieurs entreprises possèdent plusieurs filiales au sein de la ville, ce qui explique en grande partie ce taux élevé. Cette classe est également caractérisée par une redondance au niveau géométrique : plusieurs points se situent aux mêmes endroits mais ils se différencient par leur fonction. Ce résultat n’est pas anormal étant donné qu'un bâtiment peut avoir plusieurs fonctions.

Une attention particulière doit être accordée à l’exploitation de la couche *Landuse* car elle contient 33 intersections d'entités qui se traduisent par des zones « *multi-affectées* ».

Un autre cas où il faut être attentif se situe au niveau des quartiers. Deux quartiers non contigus portent le même nom, et notre connaissance du terrain n’est pas suffisante pour pouvoir identifier la localisation réelle du quartier associé au nom redondant.

Une erreur topologique a été mise en évidence au niveau des courbes de niveau sous la forme de trois auto-intersections. Celles-ci ne sont cependant pas préjudiciables à la bonne exploitation des données.

De manière pratique, le nombre d'objets par classe est déterminé dans *PgAdmin III*.

Le nombre de points par objet est évalué grâce à la requête SQL suivante, qui classe ces objets par nombre de points croissant :

```
select npoints(the_geom)
from nomDeLaClasse
order by npoints(the_geom) asc
```

Les intersections sont identifiées via la commande :

```
select *, overlaps(b1.the_geom,b2.the_geom)
from nomDeClasses b2, nomDeClasses b1
where b1.identifiant < b2. identifiant and
overlaps(b1.the_geom,b2.the_geom)='t'
```

La fonction `overlaps` permet de connaître précisément les entités qui se chevauchent. La fonction `intersects` ne nous est d'aucune utilité car la simple contiguïté y est déjà considérée comme intersection ; son utilisation n'est donc pas acceptable.

Le dénombrement des objets composés de plusieurs géométries, les multi-géométries, est réalisé grâce à la requête :

```
select *, numgeometries(the_geom)
from public."quartier polygon"
where numgeometries(the_geom) > 1
```

Les redondances géométriques sont évaluées à l'aide de la commande :

```
select *
from classe b1, classe b2
where equals(b1.the_geom, b2.the_geom) and
b1.identifiant<b2. identifiant
```

Les résultats obtenus sont cependant erronés lorsqu'il y a plus de deux entités ayant exactement la même localisation car l'algorithme ci-dessus recherche les égalités en comparant les entités deux à deux. Il faut donc en tenir compte lors du comptage des doublons.

Le comptage des redondances au niveau attributaire est effectué à l'aide de la requête suivante (exemple pour la couche *Functions*) :

```
select *
from functions b1, functions b2
where
b1.f_category=b2.f_category
and b1.f_codename=b2.f_codename
and b1.f_name=b2.f_name
and b1.f_id<b2.f_id
order by b1.f_id asc
```

Cet algorithme présente le même défaut que celui du comptage des doublons géométriques.

La recherche des parfaites redondances est menée à l'aide de la requête précédente mais en ajoutant la fonction `equals(geometry, geometry)` dans la clause `where`.

La fermeture des polygones est vérifiée grâce à la requête suivante :

```
select isclosed(the_geom)
from classe
```

Le géoréférencement ne peut pas être évalué sur base de l'algorithmique. La seule manière de l'évaluer est de se baser sur une analyse visuelle et les résultats sont donc arbitraires. La couche contenant les courbes de niveau *Contourlines*, n'a pu être corrigée et une translation d'une soixantaine de mètres vers l'ouest persiste. Il faudra en tenir compte lors de la création du Modèle Numérique de Terrain (MNT).

Il n'est pas possible de donner un indice de qualité caractérisant la précision des géométries des objets. Nous pouvons uniquement constater visuellement que certains contours d'objets ne correspondent pas exactement aux contours des objets homologues sur l'image aérienne. Nous ne pouvons pas non plus fournir une quantification de l'inexactitude de positionnement des objets par manque d'algorithmes adéquats.

Au niveau attributaire, il n'est pas possible de vérifier l'exactitude des valeurs d'attributs par suite de notre faible connaissance du terrain.

L'ensemble des résultats de ces requêtes de validation est repris à la figure 36.

D. DEMONCEAU – Prototypage d'un SIG urbain pour la gestion des risques naturels

Figure 36. Résultats de la validation des données opérationnelles

	Nombre	Géométrie	Nombre de points composant les géométries			Intersections	Nbre de multi-géométries	Auto-intersections	Redondances (%)		Parfaites redondances	Polygones non-fermés	Notes
			min	moy	max				Géométriques	Attributaires			
bridgesandculverts	140	Point	1	1	1	0	0	0	0	0	0	0	
buildings	43236	Polygone	4	5	21	1216	0	0	0		0	1	
contourlines	557	Ligne	2	364	4385	0	0	3	0		0	2	+/- 60m de décalage vers l'ouest ; 5 m d'équidistance
criticalaresandtroublespots	10	Point	1	1	1	0	0	0	0	0	0	3	
functions	969	Point	1	1	1	0	0	0	6,09	10,01	0	4	
landuse	169	Polygone	4	426	11957	33	0	0	0		0	5	8 trous
parkings	51	Polygone	4	9	26	0	0	0	0		0	6	
quarters	81	Polygone	4	6	10	0	0	0	0	1,23	0	7	
rivers	107	Ligne	2	170	2186	2	0	0	0		0	8	
roads	2145	Ligne	2	15	156	1	0	0	0		0	9	
subdivisions	3	Polygone	999	1426	1697	0	0	0	0	0	0	10	
town	1	Polygone	5	5	5	0	0	0	0	0	0	11	
villages	8	Polygone	6	55	187	0	0	0	0	0	0	12	

6.3 Accès pour une application SIG

L'accès à la base *PostgreSQL/PostGIS* est possible via *PgAdmin III*. Ce qui permet d'accéder à l'entièreté des données présentes dans la base et de manipuler celles-ci à l'aide de l'ensemble des fonctionnalités offertes par *PostgreSQL/PostGIS*. Cependant, cette méthode ne permet pas la visualisation directe des résultats. Cette visualisation n'est possible qu'ultérieurement par connexion à la base de données via un SIG-logiciel.

Figure 37. Interface de connexion à la base à partir de PgAdmin III

(Source : d'après POSTGIS 2010)

L'interface de connexion à partir de *PgAdmin III* vers la base est illustrée à la figure 37. Il faut nommer la connexion, spécifier l'adresse *ip* du serveur, le nom de la base en question, le port d'écoute, et il faut s'authentifier avec un nom d'utilisateur et un mot de passe. L'extraction des données est réalisée à l'aide de requêtes SQL et le temps d'accès dépend de la manière dont sont structurées les données dans les tables et aussi de la manière de poser les requêtes.

Une alternative pour accéder aux données est de passer par un SIG-logiciel qui a l'avantage de posséder un *viewer* permettant la visualisation des données géométriques. Presque tous ces logiciels disposent de fonctionnalités de base de données et permettent la connexion à distance à n'importe quelle base. Dans ce cas, une interface similaire à celle de *PgAdmin III* est présentée à l'utilisateur et les mêmes paramètres doivent y être insérés pour récupérer les couches de données.

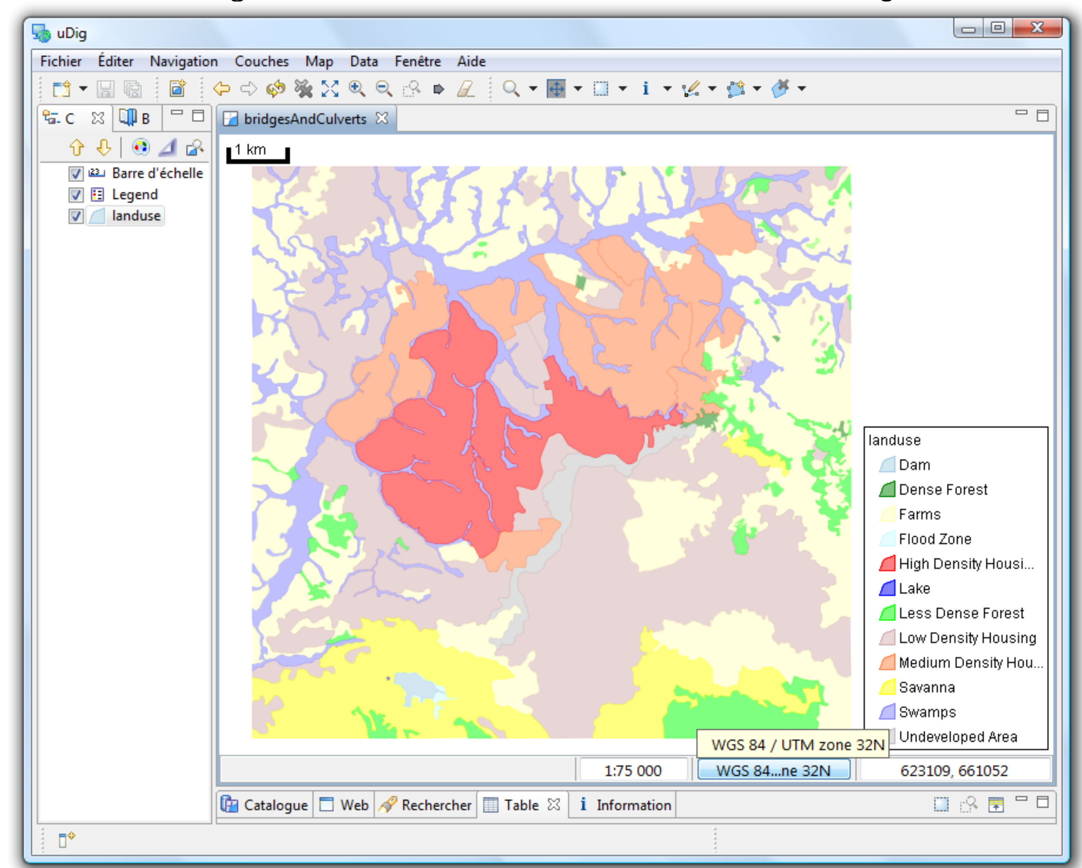
A l'exception d'*uDig*, tous ces logiciels possèdent un menu permettant de réaliser, sur les données, des requêtes spatiales et attributaires SQL.

7 Exploitation du prototype

7.1 Visualisation

Nous avons déjà parlé en section 6.3. de la visualisation des données par les *viewers* de SIG-logiciels. De manière concrète, les données brutes peuvent être visualisées comme montré à la figure 38 où la couche *Landuse* est importée dans *uDig*.

Figure 38. Visualisation de la couche landuse dans *uDig*

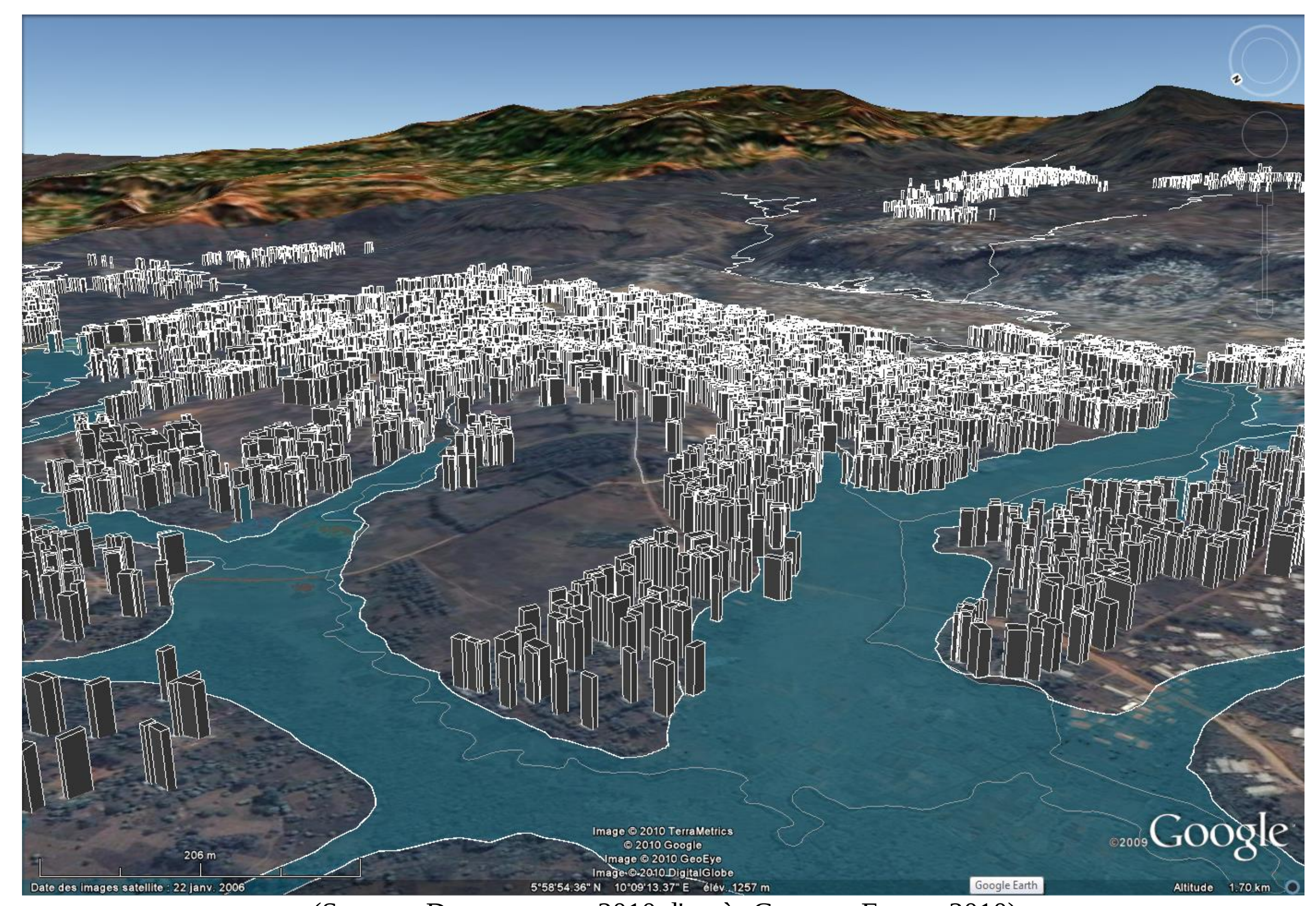


(Source : d'après UDIG 2010)

Cependant une présentation à l'aide de *viewers* ne nous semble pas présenter un caractère assez persuasif pour pouvoir être utilisée par NYAMBOD dans le but, par exemple, de convaincre les responsables de l'aménagement du territoire de la ville de Bamenda de l'utilité et de l'apport positif d'un SIG dans la gestion territoriale. Nous pensons qu'il est possible de renforcer le caractère convaincant des résultats en important les données dans un environnement tridimensionnel tel que celui de *Google Earth*. Il existe bien sûr d'autres logiciels de globe virtuel tel que *Bing Maps 3d*, *World wind* mais notre choix s'est porté sur ce logiciel gratuit parce qu'il est le plus utilisé par le grand public. Avec ce type de logiciel, il est possible de donner au spectateur l'impression qu'il se déplace réellement dans la zone d'étude.

A titre d'exemple, nous avons importé les couches contenant les zones humides, le réseau hydrographique et les bâtiments (figure 39).

Figure 39. Exemple de présentation des résultats finaux



(Source : DEMONCEAU, 2010 d'après GOOGLE EARTH 2010)

Les polygones représentant les bâtiments ont été extrudés et leur hauteur a été exagérée (50m) pour une meilleure illustration.

L'importation des *shp* dans *Google Earth* requiert un changement de coordonnées et de format des données :

- les données en format *shp* sont référencées dans la projection UTM Zone 32 N à partir du datum WGS84 et, pour les insérer dans *Google Earth*, il faut les convertir en coordonnées géodésiques WGS84 ;
- les coordonnées des entités géographiques sont disponibles en format WKT et sont accessibles sous *OpenJump*. Elles sont extraites et converties en coordonnées géodésiques. La méthodologie pour effectuer la conversion de coordonnées est explicitement décrite dans DUTCH (2009) et les formules intermédiaires sont détaillées en annexe 2 ;
- pour que ces nouvelles coordonnées soient interprétables par *Google Earth*, nous avons opté pour le format courant d'importation de données *Keyhole Markup Language* (KML), langage à balises similaire à XML.

7.2 Enrichissement de données

La base de données possède une architecture statique mais son contenu, par contre, peut faire l'objet de mises à jour, d'ajout de nouvelles entités ou de suppressions.

La modification d'une donnée attributaire est simple lorsqu'elle est réalisée à distance sous *PgAdmin III*. Il suffit de s'y connecter et de changer les valeurs en question. Si plusieurs données sont sujettes à des modifications, cette technique peut devenir grande consommatrice de temps. Il faut, dans ce cas, recourir à une requête SQL de type :

```
UPDATE class
SETchamp_x   =    nouvelle_valeur_x,
   champ_y   =    nouvelle_valeur_y,
   champ_z   =    nouvelle_valeur_z
```

Il est possible de rajouter des clauses `where` pour sélectionner les lignes concernées par les changements.

Lorsque des modifications de géométries sont à réaliser, l'introduction manuelle des requêtes SQL n'est plus réalisable suite à la complexité de la description des géométries en WKB. *Qgis* permet de modifier les géométries et de réaliser la mise à jour dans la base. Il constitue donc une alternative simple pour cette tâche.

L'ajout d'une donnée, ou plus précisément d'un ou de plusieurs tuples, au sein d'une table de la base, se réalise grâce à la fonction SQL `insert` :

```
insert into nomDeLaTable (colonne_x, colonne_y,
        colonne_z, …)
select (colonne_x, colonne_y, colonne_z, …)
from nomDeLaTableSource
where condition
```

Cela nécessite la création préalable d'une table provisoire (ici `nomDeLaTableSource`) dans laquelle sont temporairement stockées les données à rajouter avant d'en être extraites et intégrées dans la table à compléter (ici `nomDeLaTable`). Il existe, bien entendu, d'autres méthodes pour enrichir une table avec de nouvelles données mais nous présentons ici uniquement celle qui nous paraît la plus simple d'utilisation. Le *plugin* « *PostGIS Shapefile and DBF loader* » de *PostGIS* ne permet pas d'ajouter des entités dans une table déjà existante. Ce *plugin* intègre la fonction de création de table avant son peuplement.

L'ajout de nouvelles couches de données géométriques est possible à l'aide du *plugin* de *PostGIS* et la méthodologie est la même que lors de la création/peuplement des tables (voir 6.1.3).

La suppression de données est simple sous *PgAdmin III*. Il suffit simplement d'ouvrir la table et d'y supprimer les données non-voulues. La possibilité de supprimer plusieurs lignes à la fois facilite grandement la tâche.

L'enrichissement d'une base de données peut également se faire en ajoutant à une couche, qui est en mode vecteur, son équivalent image qui est plus adapté pour certaines applications. C'est par exemple le cas de la couche de courbes de niveau à partir de laquelle on peut, par interpolation, créer un MNT. L'image de sortie peut être intégrée au sein de la base afin que n'importe quel futur client puisse y accéder. Pour ce faire, il faut créer un nouvelle table contenant un champ de type OID (voir la sous-section 6.1.3.3. relative au chargement de l'image aérienne au sein de la base) et la peupler en transformant les octets de l'image en flux d'octets. L'ajout d'une nouvelle classe d'entités au sein de la base ne change en rien le modèle conceptuel qui reste statique. La nouvelle image n'aura donc qu'une existence indépendante de toutes les autres classes du MCD.

7.3 Exportation de données

Les données sont stockées sous format « *texte* » dans les tables de la base de données. Le format d'exportation des données dans *PgAdmin III* est soit le format *csv*, soit le format *txt*, mais pas le format *shp*, ce qui est regrettable étant donné l'universalité de ce format dans le monde de l'information géographique numérique. Il est cependant possible d'exporter les données de la base en *shp* à l'aide de certains SIG-logiciels, *Qgis* et *OpenJump*, entre autres. Les fichiers de sortie sont constitués du *main file*, du fichier *.shx* et du fichier *dbf*. *Qgis* produit également un fichier *prj*.

Cette étape est très intéressante pour l'utilisateur du prototype dans le sens où il peut enregistrer sur son disque dur les données contenues dans la base, et éventuellement traitées, sous forme de fichiers *shp*. Il peut, à partir de là, intégrer ces données dans ses applications et *« en faire ce qu'il veut »*. La mise à jour de la base de données par ajout de nouveaux objets récents peut se faire comme expliqué dans la section 7.2.

8 Conclusion

8.1 Évaluation

Ce travail de recherche/développement, a consisté en la conception d'un prototype SIG pour la ville de Bamenda pour les risques naturels, avec les données acquises modélisées conformément à la norme ISO 19110 et en utilisant les produits *Open Source* disponibles.

Dans le **premier chapitre,** nous avons mis en évidence le fait que les catastrophes naturelles, de quelque nature qu'elles soient, sévissent encore fréquemment dans les villes, notamment dans les pays en voie de développement. La quantité d'informations à traiter pour la gestion des risques liés à ces catastrophes est telle qu'il faut impérativement utiliser un outil SIG pour que leur analyse et leur gestion soient efficaces. Nous avons également, au travers de quelques exemples, montré quels outils y ont été mis en place ainsi que les résultats qui ont été obtenus grâce à leur exploitation.

Dans le **chapitre 2**, nous avons présenté l'objectif de la recherche. Ce mémoire a été réalisé dans le cadre d'une collaboration avec NYAMBOD, un étudiant camerounais de la VUB, et avec une répartition des tâches bien précise : nous développons un outil qu'il exploitera.

Cette collaboration a présenté des aspects positifs : les travaux de fin d'études correspondants étant très imbriqués, nous avons dû développer notre côté organisationnel, systématiser les échanges d'information, planifier les tâches de chacun de nous et respecter les *deadlines* fixées afin de permettre à NYAMBOD d'utiliser l'outil à développer. En outre, le fait qu'il soit anglophone nous a amenés à communiquer avec lui en anglais et à développer notre expression orale dans une langue autre que notre langue maternelle afin de réduire les risques de confusion et d'erreur d'interprétation.

Mais cette collaboration a également eu des aspects négatifs. Nous avons constamment été dépendants de NYAMBOD. C'est lui qui devait nous fournir les données et il a mis à jour, de sa propre initiative, certaines d'entre elles sur base de sa propre connaissance du terrain. Nous avons également rencontré de nombreuses difficultés pour obtenir des renseignements complémentaires fiables sur ces données. Le caractère erroné de certaines données reçues n'a pu être détecté que tardivement, avec les pertes de temps que cela entraîne. Ainsi, certaines métadonnées reçues telles que le datum et le système de projection étaient erronées, ce dont nous ne nous sommes rendu compte que lors de l'analyse de résultats aberrants obtenus en les exploitant. Les données fournies par l'étudiant ont dû faire l'objet d'une analyse approfondie afin de contrôler leur qualité. Le nom du satellite ayant enregistré l'image aérienne était incorrect. Certaines sources n'ont pu être retrouvées telles que, par exemple, l'origine des taux d'accroissement de la population utilisés pour calculer le nombre d'habitants à des dates postérieures au dernier recensement. Les données obtenues étaient également difficilement compréhensibles en raison de notre manque de connaissance du terrain et de l'absence d'explication préalable de la part de NYAMBOD. En outre, pour répondre à chacune de nos questions,

il était obligé de recontacter l'administration de la ville de Bamenda. Le fait qu'il y ait eu un intermédiaire entre nous et les services locaux possédant les données a compliqué notre situation et a occasionné de nombreuses pertes de temps. Tout cela nous a fortement ralentis dans l'avancement de notre travail.

C'est également dans ce chapitre qu'ont été expliquées les raisons pour lesquelles il a été décidé de faire porter notre travail sur le développement d'un prototype SIG pour la gestion des risques naturels de la ville de Bamenda. Nous avons restreint le nombre de risques pris en compte aux seuls glissements de terrain et inondations pour que la quantité de travail à réaliser soit ramenée à celle attendue dans le cadre normal d'un travail de fin d'études.

Le **chapitre 3** a été consacré aux démarches entreprises par NYAMBOD et aux travaux qu'il a réalisés en préparation du nôtre, et plus particulièrement à ceux relatifs à l'acquisition des données.

L'hypothèse de notre travail y est énoncée : *Avec les données existantes modélisées suivant la norme ISO 19110 et les logiciels Open Source disponibles, il est possible d'implémenter un prototype SIG, utile à gérer certains des problèmes urbains de la ville de Bamenda tels qu'identifiés par l'enquête de NYAMBOD*, et ce, sous la *contrainte de minimisation des coûts*.

L'analyse et la préparation des données ont été traitées au **chapitre 4**. Le grand nombre de classes différentes aurait permis de couvrir la plupart des domaines relatifs aux risques naturels et à la vulnérabilité des infrastructures pouvant entrainer des pertes tant humaines que matérielles. L'analyse a mis en évidence un grand nombre de défauts dans ces classes, défauts auxquels nous avons dû remédier en réalisant un certain nombre de traitements d'optimisation. Certains de ceux-ci constituaient des traitements originaux, notamment celui utilisé pour réaliser la correction de géométrie du lac et du réservoir de la ville et ceux portant sur la couche relative aux données de population. Cette couche a subi une série de traitements afin de permettre de déterminer la taille moyenne des ménages sur base du nombre d'habitations. Cependant, ces données n'ont pu être intégrées dans la couche de bâtiments qu'après création du modèle logique de données, au chapitre 5, car c'est seulement à partir de ce stade qu'il a été possible d'identifier les habitations. D'autres traitements ont porté sur : les noms de routes qui ont été transférés d'une couche à géométries médiocres à une couche où le réseau routier est plus dense et donc plus pertinent, les entités parkings initialement ponctuelles qui ont été converties en entités zonales, ainsi que la couche de ponts qui a été complétée par une série d'opérations. La dernière optimisation de données a consisté en une recherche des incohérences (fautes d'orthographe, fautes de frappe …) au niveau des noms dans les tables attributaires associées aux données.

Cependant, tous les traitements d'optimisation envisagés n'ont pas abouti à un résultat probant. Ainsi, une tentative de suppression des intersections entre bâtiments a échoué car elle créait plus de défauts topologiques qu'elle n'en supprimait. La correction de ces nouvelles erreurs aurait été trop fastidieuse car nous ne disposions pas d'un logiciel prenant en charge les fonctionnalités requises.

La dernière tâche de cette première étape du développement du prototype SIG a été celle du renommage de l'ensemble des données opérationnelles.

Ce chapitre s'est clôturé par la présentation de la méthodologie suivie pour créer le catalogue des données opérationnelles. Nous avons étoffé la norme utilisée (ISO 19110) en ajoutant des diagrammes UML permettant de mieux illustrer les associations entre les données. Le formalisme UML a fait

l'objet d'un ajout d'un pictogramme évocateur de la primitive géométrique des objets contenus dans les classes.

La modélisation conceptuelle de données a été décrite dans le **chapitre 5**. Cette modélisation a permis de décrire la structure du prototype. Le développement du modèle logique, qui est l'étape suivante du développement du prototype SIG, a consisté à réaliser une série de traitements pour rendre possibles les liens identifiés au niveau du modèle conceptuel de données. Nous n'avions pas sous la main un logiciel équipé d'un AGL permettant de réaliser facilement le passage du modèle conceptuel au modèle logique et nous avions donc dû réaliser ce passage manuellement à l'aide de SIG-logiciels.

Un traitement original a été effectué en vue de rendre possible les liens entre les bâtiments et les fonctions de bâtiments et ceci dans le but de savoir « quels bâtiments possèdent quelle(s) fonction(s) ». Nous avons déterminé, par une approche statistique, une distance tampon pour chacune des classes d'affectation et tout bâtiment situé à une distance inférieure à cette distance tampon de l'entité ponctuelle *Functions* a été rattaché à la classe de fonction en question.

Le **chapitre 6** a abordé le volet de l'implémentation physique du prototype. Cette dernière à consisté à créer la base de données sur le poste serveur se situant dans le bâtiment B5 et plus précisément à l'unité de Géomatique de l'ULg, à créer les tables et à les peupler avec les données modélisées. L'outil utilisé pour réaliser ces deux dernières tâches est le récent *plugin* de *PostGIS*. L'image aérienne a été intégrée au sein de la base sous la forme d'un BLOB. L'intégrité de la base a été assurée à l'aide de la définition de clés primaires et étrangères pour chacune des tables. L'implémentation physique s'est clôturée par la validation des données. Lors de celle-ci, il ne nous a pas été possible de quantifier chacun des paramètres de recherche énoncés. Ainsi, nous ne disposions pas d'une méthode permettant l'évaluation chiffrée de la précision des géométries des objets digitalisés par rapport à leurs géométries réelles représentées sur l'image aérienne. Le même problème a été rencontré pour l'évaluation de l'exactitude de position des objets ainsi que l'exactitude des attributs.

Le **chapitre 7** a été consacré à l'exploitation du prototype SIG bien que celle-ci ne soit plus du ressort du concepteur mais de celui de NYAMBOD qui, en tant que client, intègrera ces données dans un modèle de risque. L'exploitation du prototype y a été abordée sous trois aspects : la visualisation, l'enrichissement et l'exportation des données. Une manière originale de visualiser les données, traitées ou non après extraction de la base, a été développée en vue de rendre la présentation des résultats plus convaincante. Elle consiste à les intégrer dans un environnement tridimensionnel comme nous l'avions fait avec *Google Earth*.

Nous pouvons donc conclure que l'hypothèse de base de notre travail a été rencontrée : il a été possible avec les données numériques disponibles modélisées suivant la norme ISO 19110 de réaliser un prototype SIG utile pour gérer certains des problèmes urbains tels qu'identifiés dans les enquêtes de NYAMBOD. Et cela a été fait en respectant la contrainte de minimisation des coûts : les seules dépenses effectuées l'ont été pour l'acquisition des données. Aucune somme n'a été dépensée pendant la phase de conception proprement dite car les outils *Open Source* disponibles nous ont permis de réaliser les tâches souhaitées.

Nous avons également démontré que l'outil que nous avons développé est accessible à partir d'un poste externe, qu'il est possible d'interagir avec lui, et que les données sont extractibles, manipulables, convertibles et intégrables dans le cas d'une application SIG.

Les utilisations de notre prototype qui seront faites par NYAMBOD ainsi que l'illustration des résultats sous forme de cartes seront présentées dans le mémoire de celui-ci.

8.2 Perspectives

Le prototype a été développé pour la gestion des risques naturels de la ville de Bamenda. Nous pourrions appliquer une méthodologie de conception similaire pour une autre ville d'un pays en voie de développement en utilisant des produits *OpenAccess*, après évaluation du matériel informatique disponible, et moyennant une collecte adéquate et ciblée des données.

Nous avons développé ce prototype de SIG pour qu'il serve de support technique à NYAMBOD pour le développement des travaux de son mémoire. D'autres personnes, voire même les administrations locales de Bamenda, pourraient l'exploiter pour autant qu'elles trouvent un nouveau poste d'hébergement autre que le serveur du B5 et qu'elles prennent leurs responsabilités quant à la qualité des données. Car en tant que concepteur de ce prototype et dans le cadre de la collaboration, notre rôle se limitait à la conception de l'outil demandé par le client avec les données disponibles. Nous ne sommes donc responsables, ni de la qualité des données brutes acquises, ni de l'exploitation qui pourra en être faite.

Avec une collecte ciblée de données supplémentaires et surtout avec des métadonnées bien plus complètes, il serait intéressant d'étudier la manière d'extrapoler le prototype vers un SIG complet. La fiabilité du produit et ses possibilités d'utilisation seraient considérablement augmentées.

9 Bibliographie

ACHO-CHI 1998. Human interference and environmental instability: addressing the environmental consequences of rapid urban growth in Bamenda, Cameroon, *Environment and Urbanization*, 10 (2), pp. 161-174

AUBE G., BONN F., BENIE GOZE B. & MÜLLER-POITEVIEN C. 2006. Un système géomatique de préparation aux interventions d'urgence dans le bassin du Mékong, *Revue internationale de Géomatique*, **16**, pp. 301-319

AYRAL P.-A., SAUVAGNARGUES-LESAGE S. & DUSERRE G. 2001. *L'étude des risques naturels à travers un outil commun : le Système d'information géographiques*. Centre de recherche sur les Risques et les Crises (http://www.crc.mines-paristech.fr/geri/Geri2001/actesSirnat/S2_ayral.pdf), consultation le 8 février 2010

BALRAM S.& DRAGICEVIC S. 2006. Modeling Collaborative GIS Processes Using Soft Systems Theory, *UML and Object Oriented Design*, Transactions in GIS, **10**, pp. 199-218

BAZOUN J. 2008. Catastrophes naturelles et applications des SIG au Burkina Faso, *Workshop international sur « l'outil Spatial au service de la gestion des catastrophes naturelles et des situations d'Urgences en Afrique: Aspects Techniques, Organisationnels et juridiques »*, Rabat 2008, 30 p.

BEGINNING SPATIAL. Site internet (http://beginningspatial.com/common_srids), consultation le 15 mai 2010.

BILLOT C., DURLER S. & WIDMER B. 2003. Élaboration d'un outil d'estimation qualitative du risque naturel à moyenne échelle, *SIRNAT-JPRN*, Laboratoire des Systèmes d'Information Géographique, École Polytechnique Fédérale de Lausanne, 7 p.

CALKINS H. 1997. *Local Government GIS Demonstration Grant*

CASADEI M., DIETRICH W. E. & MILLER N. 2003, Testing a model for predicting the timing and location of shallow landslide initiation in soil-mantled landscapes, *Earth surface processes landforms*, **28**, pp. 925-950

COLLIGNON A.-S. 2004, Catalogue de données des objets de fontainerie de la Compagnie Intercommunale Liégeoise des Eaux d'après la norme ISO 19110, Mémoire de fin d'études en Sciences Géographiques, Université de Liège, 151 p.

DOI D. D., THANG N. V., GIANG T. B., TUAN V. A. & TOAN N. Q. 2009. Developing Vietnam's Cadastral Data Standards based on ISO 19100, *7th FIG Regional Conference*, Hanoi 2009, 17 p.

DONNAY J.-P. 1997. *Consignes de rédaction des mémoires et TFE en Géomatique et Géométrologie*, notes de cours, Université de Liège, Faculté des Sciences, inédit, 71 p.

DONNAY J.-P. 2007a. *Analyse spatiale*, notes de cours, Université de Liège, Département de Géomatique, inédit, 292 p.

DONNAY J.-P. 2007b. *Cartographie mathématique*, notes de cours, Université de Liège, Département de Géomatique, inédit, 230 p.

DONNAY J.-P. 2009. *Système d'Information Géographique*, notes de cours, Université de Liège, Département de Géomatique, inédit, 554 p.

DUCUING J.-L. 2003. L'apport des nouvelles technologies dans la gestion des risques naturels : SIG, internet,

modélisation et technologies spatiales, *SIRNAT-JPRN Orléans*, 6 p.

DUTCH S. 2009, *Converting UTM to Latitude and Longitude (Or Vice Versa).* Site non-officiel de l'Université du Wisconsin Green Bay (http://www.uwgb.edu/dutchs/usefuldata/utmformulas.htm), consultation le 9 avril 2010,

FOURNEAUX D. 2005. *Portail interopérable d'informations géographiques*, Mémoire de fin d'études en Sciences Géographiques, Université de Liège

HUXHOLD W. E. 1991. *An Introduction to Urban Geographic Information Systems*, Oxford : Oxford University Press, 376 p.

INSTITUT GEOGRAPHIQUE NATIONAL FRANÇAIS (IGNF) 2002, Base de données topographique/Agglomération – Spécifications de contenu v1.2, 148 p.

INTERNATIONAL ORGANIZATION FOR STANDARDISATION ISO/DIS 2001SO/TC 211 19110. Geographic information – Metodology for feature cataloguing, 2001, 60 p.

LAPLANCHE F. 2001. *Modélisation conceptuelle de projet SIG avec UML*, Mémoire de fin d'études en Sciences Géographiques, Université de Liège

LAPLANCHE F. 2002. Conception de projet SIG avec UML, *Bulletin de la Société géographique de Liège*, 42, pp. 19-25

LAPLANCHE F. 2006, *Environnement de conception de bases de données spatiales sur internet*, Thèse de troisième cycle en Sciences Géographiques, Université de Liège

LAURINI R. 2001. *Information Systems for Urban Planning*, London : Taylor & Francis, London, 370 p.

LEDENT A. 2004, *Applicatif SIG au service Patrimoine de la CILE*, Mémoire de fin d'études en Sciences Géographiques, Université de Liège

MAKON MA PONDI 2009. *Effondrement de la route à Bamenda: Impératif la maintenance !*, Cameroon Tribune (http://www.cameroon-info.net/reactions/@,25292,7,effondrement-de-la-route-a-bamenda-imperatif-la-maintenance.html), consultation le 20 mai 2010

MALET J.-P., THIERY Y., MAQUAIRE O. & PUISSANT A. 2006. Analyse spatiale, évaluation et cartographie du risque « glissement de terrain », *Revue internationale de Géomatique*, **16**, pp. 499-525

MYSQL 2010, *Well-Known Binary (WKB) Format*, (http://dev.mysql.com/doc/refman/5.0/en/gis-wkb-format.html), consultation le 15 mai 2010

NATIONAL GEOSPATIAL-INTELLIGENCE AGENCY 2008. *Africa Datums Table* (http://earth-info.nga.mil/GandG/coordsys/onlinedatum/CountryAfricaTable.html), consultation le 5 avril 2010

NJOG M. N. 2009. *Wum – Une ville à la recherche de la modernité*, AllAfrica.com (http://fr.allafrica.com/stories/200908060836.html), consultation le 20 mai 2010

NDULA Y. & NDEFRU N. 2009. *Landslide Blocks Bamenda Station Hill Road*, AllAfrica.com (http://allafrica.com/stories/200908050315.html), consultation le 21 mai 2010

NSEUMI LEA F. 2009. *Glissements de terrains – La commission de précention des catastrophes portée disparue*, Cameroun-online (http://www.cameroun-online.com/actualite,actu-11690.html), consultation le 20 mai 2010

NYAMBOD E. M. 2010. Environmental Consequences of Rapid Urbanization: Bamenda City, *Cameroon, Journal of Environmental Protection*, **1**, pp. 15-23

OPENJUMP 2010. OpenJump : An open source Geographic Information System written in Java (http://www.openjump.org)

PENNOBER G., GIRAUDET E., GIRAUDET J., MADEC-CUQ V., GOURMELON F., SIMAO DA SILVA A. & CAMPREDON P. 2005. Planification côtière en Afrique de l'Ouest. Retour d'expérience SIG en Guinée-Bissau, *Norois,* **3**, pp. 196-236

PETIT F. 2007. *Introduction à la Géomorphologie*, notes de cours, Université de Liège, Département de Géomatique, inédit, 137 p.

PETIT F. 2009. *Hydrologie appliquée*, notes de cours, Université de Liège, Département de Géomatique, inédit, 152 p.

PIECHOCKI L. 2007. *UML, le langage de modélisation objet unifié*. Developpez.com, Club des développeurs (ftp://ftp-developpez.com/laurent-piechocki/uml/tutoriel/lp/cours/coursUml.pdf), consultation le 19 février 2010

PORNON H. 1992. *SIG pour petites communes.* Paris : Edition du STU Ministère de l'Équipement, 127 p.

POSTGRESQL 2010. PostgreSQL: The world's most advanced open source database. Site internet (http://www.postgresql.org)

PRIOU D., LE GALLIC J.-M., PELLE S. & RICHARD D. 2004. *UML, Java et les données géographiques : notions de base*, collection ENSG-IGN, 1ère édition. Paris : Hermès - Lavoisier, 470 p.

QGIS 2010. Quantum GIS – The new Direction of Open Source (http://www.qgis.org)

SANGA D. & DOSSO B. 2007. L'utilisation des SIG dans les Instituts/Bureaux nationaux de statistique africains, *Le Journal statistique africain*, **5**, pp. 161-181

SIGAFRIQUE 2007. *SIGAfrique : Réseau africain d'information géologique pour le développement durable.* Site internet consulté le 20 mai 2010 (http://www.sigafrique.net/Presentation.aspx)

SYBASE 2010. *Worldwide [change] : Managing, Analyzing and Mobilizing Information*. Manuel d'utilisation. Site internet (http://www.sybase.com/), consultation le 2 juin 2010

TONYE E. & AKONO A. 2002. Two approaches of textural classification for the location of the flows of lava on the volcanic site of Mountain Cameroon, *Geocarto International*, **17**, pp. 35-46

WARNANT R. 2008, *Théorie des erreurs, notes de cours*, notes de cours, Université de Liège, Faculté des Sciences, inédit, 31 p.

WIKIPEDIA 2010. *Description of the town of Bamenda* (http://en.wikipedia.org/wiki/Bamenda), consultation le 20 mai 2010

WOLPER P. 2008. *Bases de données (organisation générale)*, notes de cours, Université de Liège, Faculté des Sciences Appliquées, inédit, 344 p.

WORBOYS M.F. & DUCKHAM M. 2001. *GIS - A computing Perspective, Second edition*, CRC Press, 431 p.

10 Annexes

10.1 Dictionnaire des données brutes

1	Artirary road buffer.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	1
Bufferdis	Entier	55	1
Area	Entier	3376639	1

2	Bamenda satellite towns.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0 – 108	117
Code	Entier	1	1
		2	1
		3	4
		5	111
Name	Liste*		
Province	Liste	NORTH – WEST	115
		NORTH WEST	1
		OUEST	1
Division	Liste	BAMBOUTOS	1
		BOYO	1
		KADEY	2
		MEZAM	96
		MOMO	15
		NGOKETUNDJIA	2
Subdiv	Liste	BABADJOU	1
		BAFUT	19
		BALI	4
		BALIKUMBAT	1
		BAMENDA	15
		BATIBO	2
		BELO	1
		MBENGWI	15
		NDOP	1
		SANTA	15
		TUBAH	42
District	/		
Country	Liste	CAMEROUN	116
X_coord	Décimale	0	1
		604725 – 650657	
Y_coord	Décimale	0	1
		636583 – 680509	
Loca_code	Entier	0	29
		7	88

3	Bridges and culverts.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	25
		1	63
		2	2
Class	Liste	PONTS	63
		CULVERT WOODER	2
		CULVERTS	25
Material	/		

4	Buffer round critical junctions.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Bufferdis	Entier	200	10

5	Catchment sec general and technical.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	1
Bufferdis	Entier	2000	1

6	cbd.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	8
Type	Liste	CBD	1
		FL	2
		FLSC	2
		SSC	1
		TSC	2
Name	Liste	Central Business District	1
		Fourth Level Service Centre	2
		Fourth Level Service Centre OS	2
		Secondary Service Centre	1
		Teriary Service Centre	2

7	Contour satellite towns.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Fnode	Entier	43719 – 80696	
Tnode	Entier	43719 – 80697	
Length	Décimale	475,816 – 19274,894	
Con50sm_id	Entier	53625 – 102983	
Contour	Entier	600 – 2550	1

8	Council halls.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	2
Nom	Liste	NEW SITE	1
		OLD SITE	1

9	Courbes de niveau bamenda.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1 – 504	
Altitude	Entier	1200 – 2135	
Type	Liste	Thick	274
		Thin	283

10	Courbes de niveau bamenda_urban area.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1 – 434	
Altitude	Entier	1200 – 1865	
Type	Liste	Thick	271
		Thin	268

11	Critical areas and trouble spots.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1 – 10	
Name	Liste	City Chemist Round About	1
		Council Junction	1
		Foncha Street Junction	1
		Hospital Round About	1
		Mobil Nkwen	1
		Mrs Ngeng Junction	1
		PMI Junction	1
		Rota Nkwen	1
		Total Nkwen	1
		Vetirinary Junction	1

12	Eauip machands.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X coord	Entier	623913 – 631991	
Y coord	Entier	655856 – 667367	
Code	Entier	2	5
		3	1
		4	11
		5	3
		6	1
		7	1
		8	17
		9	1
		11	5
		12	43
Code_nom	Liste	Abattoir	1
		Aerôport	1
		Autres	34
		Cooperative agricoles	1
		Gares routiére\Agences de voyages	26
		Marché d'animaux	1
		Marché de vivres	5
		Marché du poisson	2
		Marché mixte	3
		Station service	5
		Station Service	9
Nom equipe	Liste*		
Quartier	/		
Ville	Liste	BAMENDA	80
Commune	Liste	BAMENDA URBAN COUNCIL	80
Arrondisse	Liste	BAMENDA CENTRAL	80
Departemen	Liste	MEZAM	80
Province	Liste	NORTH WEST PROVINCE	80

13	Education.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X coord	Entier	623559 – 634752	
Y coord	Entier	653320 – 664618	
Code	Entier	0	1
		1	45
		2	50
		3	33
		4	28
		5	3
		6	14
		7	1
		8	5
Nom_code	Liste	Enseignement	3
		Formation professionnelle	27
		Maternelle	1
		Maternelle/primaire	94
		Primaire	20
		Recherche	6
		Sécondaire général	28
		Université/grand école	1
Nom de l i	Texte*		
Quartier	/		
Quartier	/		
Ville	Texte	BAMENDA	101
Commune	Texte	BAMENDA URBAN COUNCIL	101
Arrondisse	Texte	BAMENDA CENTRAL	101
Departemen	Texte	MEZAM	101
Province	Texte	NORTH WEST PROVINCE	101
Owner	Liste	GOVERNMENT	38
		ISLAMIC	1
		PRIVATE	93
		SYSTEMIC	48

14	Education0 0707.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X coord	Entier	623559 – 634752	
Y coord	Entier	653320 – 664618	
Code	Entier	0	1
		1	45
		2	50
		3	33
		4	28
		5	3
		6	14
		7	1
		8	5
Nom_code	Liste	Enseignement	3
		Formation professionnelle	27
		Maternelle	7
		Maternelle/primaire	65
		Primaire	43
		Recherche	6
		Sécondaire général	28
		Université/grand école	1
Nom de l i	Liste*		
Quartier	/		
Quartier	/		
Ville	Texte	BAMENDA	101
Commune	Texte	BAMENDA URBAN COUNCIL	101
Arrondisse	Texte	BAMENDA CENTRAL	101
Departemen	Texte	MEZAM	101
Province	Texte	NORTH WEST PROVINCE	101
Owner	Liste (texte)	GOVERNMENT	38
		ISLAMIC	1
		PRIVATE	93
		SYSTEMIC	48

15	Equipe public et parapublic.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X_coord	Entiers	623958 – 633582	
Y_coord	Entiers	655964 – 663848	
Code	Entiers	1	92
		2	3
		4	2
		5	2
		6	1
		9	78
Nom_code	Texte	Administration	92
		AES_SONEL	2
		FEICOM	1
		ONG/GIC	78
		Organasitions internationale	3
		SNEC	2
Nom_de_ser	Liste*		
Quartier	/		
Quartier	/		
Ville	Texte	BAMENDA	178
Commune	Texte	BAMENDA URBAN COUNCIL	178
Arrondisse	Texte	BAMENDA CENTRAL	178
Departemen	Texte	MEZAM	178
Province	Texte	NORTH WEST PROVINCE	178

16	Equipements d'hôtellerie et touristiques.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X_coord	Entier	624947 – 632482	
Y_coord	Entier	654845 – 661394	
Code	Entier	1	5
		2	13
		3	5
		4	15
		6	1
		10	30
Nom_code	Texte	Auberges	6
		Chefferie ou lamidat	1
		H?tellerie	3
		Hôtelleire	21
		Restaurants	15
		Snack bar	23
Nom	Liste*		
Quartier	Texte	Bamenda	30
Quartier	Texte	Bamenda	30
Ville	Texte	BAMENDA	69
Commune	Texte	BAMENDA URBAN COUNCIL	69
Arrondisse	Texte	BAMENDA CENTRAL	69
Departemen	Texte	MEZAM	69
Province	Texte	NORTH WEST PROVINCE	69

17	Equipements marchandes.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X_coord	Entier	0	23
		623913 – 631991	
Y_coord	Entier	0	23
		655856 – 667367	
Code	Entier	0	23
		2	5
		3	1
		4	11
		5	3
		6	1
		7	1
		8	17
		9	1
		11	5
		12	43
Code_nom	Texte	Abattoir	1
		Aerôport	1
		Autres	34
		Cooperative agricoles	1
		Gares routiére\Agences de voyages	26
		Marché d'animaux	1
		Marché de vivres	5
		Marché du poisson	2
		Marché mixte	3
		Station Service	14
Nom_equipe	Liste*		
Quartier			
Ville	Texte	BAMENDA	88
Commune	Texte	BAMENDA URBAN COUNCIL	88
Arrondisse	Texte	BAMENDA CENTRAL	88
Departemen	Texte	MEZAM	88
Province	Texte	NORTH WEST PROVINCE	88

18	Evolution.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	5
		1	3
		2	13
		3	12
		4	7
		5	11
Period	/		
Remarks	/		

19	Evolving hierarchy of settlements.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X coord	Entier	594771 – 659935	
Y coord	Entier	605236 – 672677	
Item	Entier	1	1
		2	1
		3	1
		4	1
		5	1
		6	1
		7	1
		8	1
		9	1
		10	1
		11	1
Settlement	Texte	Bafoussam	1
		Bafut	1
		Bali	1
		Bambili	1
		Bambui	1
		Bamenda	1
		Batibo	1
		Mbengui	1
		Mbouda	1
		Ndop	1
		Santa	1
Dist from	Entier	0-80	
Road chara	Texte	TARRED	10
Pop 2000	Entier	6499-229200	

20	Finance.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X coord	Entier	626517 – 632054	
Y coord	Entier	655518 – 662688	
Code	Entier	1	16
		2	35
		3	16
		4	6
		5	1
		6	2
Nom_code	Texte	Banques	16
		Etablissement de crédit/Associations et coopératives	35
		Express Union	6
		Insurance/Assurance	3
		Insurances/assurances	13
		MoneyGram	1
		Western Union	2
Num de l i	Liste*		
Quartier	/		
Ville	Texte	BAMENDA	50
Commune	Texte	BAMENDA URBAN COUNCIL	50
Arrondisse	Texte	BAMENDA CENTRAL	50
Departemen	Texte	MEZAM	50
Province	Texte	NORTH WEST PROVINCE	50

21	Gradient analysis.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id		0	854

22	Habitatcub0 0207.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	/		
Prim_catch	Booléen	NO	13326
		YES	31690
Nurs_catch	Booléen	NO	34938
		YES	10078
Sec_catch	Booléen	NO	14624
		YES	30392

23	Health.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X coord	Entier	624898 – 631576	
Y coord	Entier	653971 – 664134	
Code	Entier	1	4
		2	33
		3	10
		4	4
		7	2
		8	6
		9	6
		10	2
Nom_code	Texte	Autres	2
		Centre de santé	9
		Clinique	19
		Dispensaires	2
		Hôpital	4
		Laboratoire	6
		Maternity	2
		Médecine traditionelle	6
		Pharmacie	17
Nom de l i	Liste*		
Quartier	/		
Quartier	/		
Ville	Entier	BAMENDA	67
Commune	Entier	BAMENDA URBAN COUNCIL	67
Arrondisse	Entier	BAMENDA CENTRAL	67
Departemen	Entier	MEZAM	67
Province	Entier	NORTH WEST PROVINCE	67
Owner	Entier	GOVERNMENT	9
		PRIVATE	50
		SYSTEMIC	8

24	Hotel de ville.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	/		

25	Houses within catchment zone.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	/		

26	Hydro bda lakes.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id		0	3

27	Hydro bda rivers.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	108
Code	Entier	1	3
		2	40
		3	65
Nom	Texte	Ayaba Stream	1
		Mugheb Stream	1
		River Mezam	3
Quartier	/		
Ville	Entier	Bamenda	108
Arrond	Entier	Bamenda-Central	108

28	Hydro_satellite towns.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Length	/		
Objectid	/		
Hydr_lin		1 – 1778	
Indicatif	Texte	10100101	6884
		10100102	5087
		10100103	10628
		10100104	17794
		10100200	175
		10100500	1176
		40100000	67
		480101001	1
Toponyme	Texte	04	1

29	Industry.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X_coord	Entier	624731 – 634287	
Y_coord	Entier	657694 – 663719	
Code	Entier	1	1
		2	6
		3	1
		4	22
		5	6
		6	15
		9	7
		10	3
		12	8
Nom_code	Texte	Autres	9
		Brasseries	2
		Chemique	1
		Entrep?rt industriel	2
		Imprimerie	5
		Mécanique/métallique	15
		Textile	1
		Transformation des Aliments	5
		Transformation des Aliments	1
		Transformation du Bois	21
		Transformation du Bois	1
		Travaux publique et bâtiments (BTP)	4
		Travaux Publiques	2
Nom_de_l_i	Liste*		
Quartier	/		
Ville		BAMENDA	69
Commune	Texte	BAMENDA URBAN COUNCIL	69
Arrondisse	Texte	BAMENDA CENTRAL	69
Departemen	Texte	MEZAM	69
Province	Texte	NORTH WEST PROVINCE	69

30	Internal urban spatial structure.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	14
Area	Double	566406,3 – 23790472,9	
Name	Texte	Inner Area	1
		Middle Area	3
		Peripheral Area	6
		Virgin Area	4

31	Land marks.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	242
Name	Liste*		
Status	Texte	Important Points	19
		Land Marks and Mounements	223

32	Landuse.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1	13
		2	28
		3	27
		4	25
		5	2
		6	49
		7	11
		8	10
		9	3
Code	Entier	1	129
		2	10
		3	29
Sub_code	Entier	0	39
		1	7
		2	45
		3	5
		4	72
Type	Texte	Dam	1
		Dense Forest	5
		Farms	58
		Flood Zone	1
		High Density Housing	1
		Lake	1
		Less Dense Forest	41
		Low Density Housing	34
		Medium Density Housing	6
		Savanna	3
		Swamps	16
		Undeveloped Area	1

33	Landuse points.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1	13
		2	28
		3	27
		4	25
		5	2
		6	49
		7	11
		8	10
		9	3
Code	Entier	1	129
		2	10
		3	29
Sub_code	Entier	0	39
		1	7
		2	45
		3	5
		4	42
Type	Texte	Dam	1
		Dense Forest	5
		Farms	58
		Flood Zone	1
		High Density Housing	1
		Lake	1
		Less Dense Forest	41
		Low Density Housing	34
		Medium Density Housing	6
		Savanna	3
		Swamps	16
		Undeveloped Area	1

34	Landuse points1207.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	214
Lu	Texte	BG	2
		CBD	2
		CC	1
		CCC	1
		CE	2
		CPG	2
		DC	12
		DH	1
		DPS	1
		GH	1
		H	1
		HD	65
		HI	1
		HI/LI/SI	1
		LD	9
		LI/SI	1
		M	9
		MCR	7
		MD	21
		MRS	4
		OSP	5
		OSP/TC	1
		PCF	1
		PPS	1
		PSS	1
		PT	1
		SA	18
		SI	2
		SPS	4
		SS	17
		ST	2
		TC	4
		TP	4
		TS	2
		TSH	5
		UN	1

35	Landuse polygons 1207.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	12

36	Landuse study area.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1	13
		2	28
		3	27
		4	31
		5	2
		6	51
		7	11
		8	10
		9	7
Code	Entier	1	135
		2	16
		3	29
Sub_code	Entier	0	45
		1	9
		2	44
		3	5
		4	77
Type	Texte	Dam	1
		Dense Forest	5
		Farms	61
		Flood Zone	1
		High Density Housing	1
		Lake	1
		Less Dense Forest	40
		Low Density Housing	34
		Medium Density Housing	7
		Savanna	5
		Swamps	23
		Undeveloped Area	1

37	Limit satillite town area.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Length	Double	97,655 – 16714,036	
Objectid	Entier	6137 – 34968	
Hydr lin	Entier	13 – 1123	
Indicatif	Texte	10100102	16
		10100103	82
		10100104	119
		10100200	4
Toponyme	/		

38	Limit urban area bamenda line.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	1
Length	Double	39603,64	1

39	Limit urban area bamenda.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1	1

40	Limite arrondissement.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	3
Nom	Texte	ARRONDISSEMENT I	1
	Texte	ARRONDISSEMENT II	1
	Texte	ARRONDISSEMENT III	1
Area m	Entier	20522519 – 52396394	

41	Line covering sat towns.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	1
Length	Entier	147305665	1

42	Nursery school travelling distance govt.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Bufferdis	Entier	500	1

43	Parking facilities.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1 – 50	
Remarks	Texte	CLANDESTINE AND OBSTRUCTS	45
		GOOD	5
Service	Texte	MOTORTAXI_OBSTRUCTION	10
		VEHICLE GOOD	5
		VEHICLE PARKING_OBSTRUCTION	35

44	Parking facilities revised.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	52
Nom_code	Texte	Council	4
		Motor Bike Park	12
		Motor Park	12
		Private Park	24
Name	Texte	Amour Mezam Express	2
		Bafut Park	1
		Bali Park	1
		Guarantee Express	4
		Illegal	15
		Illegal Park	2
		Jean Jeannot Express	3
		Jeannot Express	2
		Mankon Rural	2
		Mazi Voyage	1
		Mile 8 Mankon Park	1
		Mile Four Park	1
		Moghamo Travel Agency	1
		Mokas Travel Agency	2
		Mondial Express	2
		Nitob Park	1
		Njekwa Park	1
		Nkwen Park	1
		Ntatru Park	1
		Patience Express	1
		Santa Park	1
		Symbol of unity Express	1
		Symbol of Unity Travel Agency	1
		Vatican Express Agency	2
		Wum Park	1
Status	Texte	Council	6
		Illegal	22
		Private	24
Pos situti	/		

45	Population figures.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Xcoord	Entier	624442 – 631037	
Ycoord	Entier	656017 – 660348	
Id	Entier	1 – 14	
Viilage qu	Liste*		
Pop 87	Entier	1174 – 13262	
Pop 1992	Entier	1376 – 15545	
Pop 2005	Entier	3777 – 52187	

46	Primary catchment zone.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	1
Bufferdis	Entier	1000	1

47	Primary road buffer.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	9
Bufferdis	Entier	50	9
Area	Entier	63889 – 976321	

48	Quarters mendankwe.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	32
Name	Liste*		

49	Quarters bamenda.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0 – 54	
Nam quarte	Liste*		
Village			
Pop 87	Entier	0	55
Pop 1992	Entier	0	55
Pop 2005	Entier	0	55
Xcoord	Entier	623786 – 633223	
Ycoord	Entier	654463 – 662367	

50	Quarters mendankwe.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	32
Name	Liste*		

51	Quarters0 0707.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Liste*	0	59
		Autres	
Name	Liste*		
Village	Texte	Mankon	3
		Mbatu	1
		Mendankwe	34
		Nkwen	21

52	Quartier polygon.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Perimeter	Double	0 – 31841,288	
Area_km	Double	0 – 87,08	
Color	Entier	0	1
		1	9
		2	7
		3	6
		4	5
		5	6
		6	6
		7	5
		8	5
		9	4
		10	5
Subdiv_id	Entier	0	1
		6	58
Province	Texte	Nord_Ouest	58
Division	Texte	Mezam	58
Subdiv	Texte	Bamenda	58
Area_m	Double	0 – 8708069,157	
Quarter_na	Liste*		
Urb_rural	Texte	Rural	1
		Space	1
		Urban	12

53	Religious institution.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X_coord	Entier	0 – 633297	
Y_coord	Entier	0 – 664609	
Code	Entier	0	1
		1	156
		2	5
		3	6
Code_nom	Texte	Cimetière	7
		Eglise	156
		Mosquée	5
Nom_de_l_i	Liste*		
Quartier	/		
Ville	Texte	BAMENDA	167
Commune	Texte	BAMENDA URBAN COUNCIL	167
Arrondisse	Texte	BAMENDA	167
Departemen	Texte	MEZAM	167
Province	Texte	NORTH WEST PROVINCE	168

54	Rivers sat towns.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Length	Double	97,655 – 16714,036	
Objectid	Entier	6137 – 34968	
Hydr_lin	Entier	13 – 1123	
Indicatif	Texte	10100102	16
		10100103	82
		10100104	119
		10100200	4
Toponyme	/		

55	Road accessibility.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	0
Bufferdis	Entier	35	1
		45	1
		50	1
		55	1
Area	Entier	1493244	1
		2428636	1
		3376639	1
		4057941	1

56	Road network.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Long_km	Double	0 – 5,5	
Larg_m	Double	0 – 5539	
Priorite	Entier	0	2129
Rt_class	Texte	Piste	1158
		Voies à grande circulations	2
		Voies principales bitumée	26
		Voies principales en terre	1
		Voies secondaires bitumée	32
		Voies secondaires en terre	79
		Voies secondaires intermittente bitumée	22
		Voies secondaires intermittente en terre	809
Rt_surface	Texte	Bitumée	79
		Paved	2
		Terre	2046
Rt_etat	Texte	BAD	55
		Bon	40
		Mauvaise	1759
		Moin Mauvaise	267
Action_pro	/		
Rt_nom	/		
Quartier	/		
Ville	Texte	BAMENDA	2115
Commune	/		
Arrond	/		
Départ	/		
Province	/		

57	roads linkin the evolving hieracchy of satellite towns.shp		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Double	0 – 5,5	
Tarred	Entier	0	28
Province	Texte	NORTH – WEST	26
		OUEST	2
Road_Code	Texte	P24	2
		P18	5
		N6	2
		N4	1
		N11	2
Subdiv	Texte	BADADJOU	1
		BAFUT	2
		BAMENDA	3
		BATIBO	2
		MBENGWI	3
		MENOUA	1
		NDOP	1
		TUBAH	15
Length_km	Entier	0 – 30	
Road_class	Texte	PROVINCIAL	12
		National	16
Length_m	Entier	0 – 16587	
Division	Texte	BAMBOUTOS	1
		MEZAM	12
		MIFI	1
		MOMO	4
		NGOKETUNJIA	10
Country	Texte	CAMEROUN	28
Surface	Texte	Untarred	9

58	Roads_0770.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Long_km	Double	0 – 5,5	
Larg_m	Double	5 – 5539	
Priorite	Entier	0	2145
Rt_class	Texte	ARTILIARY	14
		PRIMARY DISTRIBUTOR	51
		SECONDARY ROADS	50
		TERTIARY ROAD	36
Rt_surface	Texte	Bitumée	80
		Paved	2
		Terre	2063
Rt_etat	Texte	Bon	47
		Mauvaise	1811
		Moin Mauvaise	287
Rt_nom	/		
Quartier	/		
Ville	Texte	BAMENDA	2133
Commune	/		

59	Secondary road buffer.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	1
Bufferdis	Entier	45	1
Area	Entier	2428636	1

60	Securité publique.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X_coord	Entier	625099 – 631242	
Y_coord	Entier	655750 – 662959	
Code	Entier	1	8
		2	14
		5	1
		8	2
		10	4
Nom_code	Texte	Camp Millitaire	1
		Gendamerie nationale	8
		Police	13
		Pompes funèbres	4
		Prison	1
		Service securité	2
Nom	Liste*		
Quartier	/		
Quartier	/		
Ville	Texte	BAMENDA	29
Commune	Texte	BAMENDA URBAN COUNCIL	29
Arrondisse	Texte	BAMENDA CENTRAL	29
Departemen	Texte	MEZAM	29
Province	Texte	NORTH WEST PROVINCE	29

61	Service public et parapublic.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X_coord	Entier	623958 – 633582	
Y_coord	Entier	655964 – 663848	
Code	Entier	0	4
		1	84
		2	3
		4	2
		5	2
		6	1
		9	78
Nom_code	Entier	Administration	87
		AES_SONEL	2
		FEICOM	1
		ONG/GIC	78
		Organasitions internationale	3
		SNEC	2
Nom_de_ser	Liste*		
Quartier	/		
Quartier	/		
Ville	Texte	BAMENDA	171
Commune	Texte	BAMENDA URBAN COUNCIL	171
Arrondisse	Texte	BAMENDA CENTRAL	171
Departemen	Texte	MEZAM	171
Province	Texte	NORTH WEST PROVINCE	171
Status	Entier	0	1
		1	171
		2	1
		3	1
Level	Texte	DIVISIONAL AND SUBDIVISIONAL	32
	Texte	PROVINCIAL	47

62	Station escarpement.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	1

63	Subdivibdapolygon0_0707.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1	1
		2	1
		3	1
Subdivbda	Texte	BAMENDA I SUBDIVISION	1
		BAMENDA II SUBDIVISION	1
		BAMENDA III SUBDIVISION	1
Division	Texte	MEZAM	3
Area	Entier	22822052	1
		24587917	1
		48905035	1

64	Subdivibdapolyline0_0707.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1	1
		2	1
		3	1

65	Subdivision line.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	1	1
		2	1
		3	1

66	Subregional context_merge.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	9
		Autres 5 – 298	
Code	Entier	1	4
		2	9
		3	6
		5	5
Name	Liste*		
Province	Texte	ADAMAOUA	1
		EXTREME - NORD	6
		NORTH - WEST	9
		OUEST	8
Division	Liste*		
Subdiv	Liste*		
District	/		
Country	Texte	CAMEROUN	24
Pop	Entier	0 – 302000	24

67	Swamps.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	2	2
		3	2
		4	9
		6	3
		7	1
		8	1
Code	Entier	1	4
		2	12
		3	2
Sub_code	Entier	0	14
		2	1
		4	3
Type	Texte	Swamps	18

68	Télécommunication.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
X_coord	Entier	624832 – 631031	
Y_coord	Entier	656743 – 664208	
Code	Entier	0	1
		2	3
		3	4
		4	6
		5	4
		6	5
		7	2
		8	17
		9	35
		10	2
Code_nom	Texte	Autres	2
		Camtel	2
		Centrale télephonique	2
		Courrier	6
		MTN	4
		Orange	5
		Presse	10
		Radio/TV	8
		Service internet	36
		Station relais autonomes	7
Nom	Liste*		
Quartier	/		
Quartier	/		
Ville	Texte	BAMENDA	81
Commune	Texte	BAMENDA URBAN COUNCIL	81
Arrondisse	Texte	BAMENDA CENTRAL	81
Departemen	Texte	MEZAM	81
Province	Texte	NORTH WEST PROVINCE	81

69	Travelling distance sec school.shp		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Bufferdis	Double	2000.0	1

70	Tertary road buffer.*shp*		
Attributs	*Type valeur*	*Valeur*	*Occurrence*
Id	Entier	0	1
Bufferdis	Entier	35	1
Area	Entier	1493244	1

10.2 Formules de conversions (x,y) UTM zone 32 N → (ϕ,λ) WGS 84

Il faut tout d'abord calculer l'arc méridional :

$$M = \frac{y}{k_0}$$

y est la coordonnée UTM du point à convertir et k_0 vaut l'échelle le long de la longitude centrale de la zone et vaut 0,9996.

Nous calculons ensuite la *Footprint Latitude*, qui est la latitude du point en considérant le méridien central comme origine, via les formules :

$$\mu = \frac{M}{a\left(1 - \frac{e^2}{4} - \frac{3e^4}{64} - \frac{5e^6}{256} \ldots\right)}$$

$$e_1 = \frac{[1-(1-e^2)^{\frac{1}{2}}]}{[1+(1-e^2)^{\frac{1}{2}}]}$$

$$fp = \mu + J_1 \sin(2\mu) + J_{2\sin}(4\mu) + J_{3\sin}(6\mu) + J_{4\sin}(8\mu)$$

e vaut dans ce cas l'excentricité de l'ellipsoïde dans un plan contenant l'axe de rotation de la terre qui vaut :

$$e = \sqrt{1 - \frac{b^2}{a^2}}$$

avec a=6378137 et b=6356752,31 pour l'ellipsoïde GRS80.
et les constantes J_1 à J_4 valent

$$J_1 = \left(\frac{3e_1}{2} - \frac{27e_1^3}{32} \ldots\right)$$

$$J_2 = \left(\frac{21e_1^2}{16} - \frac{55e_1^4}{32} \ldots\right)$$

$$J_3 = \left(\frac{151e_1^3}{96}\right)$$

$$J_4 = \left(\frac{1097e_1^4}{512}\right)$$

La latitude est calculée via :

$$lat = fp - Q_1(Q2 - Q3 + Q4)$$

Les variables intermédiaires à calculer pour déterminer longitudes et latitudes se calculent grâce à :

$$Q_1 = N_1 \frac{\tan(fp)}{R_1}$$
$$Q_2 = (\frac{D^2}{2}) Q_3 = (5 + 3T_1 + 10C_1 - 4C_1^2 - 9e'^2)\frac{D^6}{24}$$
$$Q_4 = (61 + 90T_1 + 298C_1 + 45T_1^2 - 3C_1^2 - 252e'^2)\frac{D^6}{720}$$

où les inconnues valent :

$$e'^2 = \frac{e^2}{(1-e^2)}$$
$$C_1 = e'^2 \cos^2(fp)$$
$$T_1 = \tan^2(fp)$$
$$R_1 = \tan^2(fp)$$
$$N_1 = \frac{a}{(1 - e^2 \sin^2(fp))^{\frac{1}{2}}}$$
$$D = \frac{x}{N_1 k_0}$$

La longitude géodésique peut être déterminée grâce à :

$$long = long0 - \frac{Q5 - Q6 + Q7}{\cos fp}$$

où les inconnues Q_5, Q_6 et Q_7 sont calculées via :

$$Q_5 = D$$
$$Q_6 = (1 + 2T_1 + C_1)\frac{D^3}{6}$$
$$Q_7 = (5 - 2C_1 + 28T_1 - 3C_1^2 + 8e'^2 + 24T_1^2)\frac{D^5}{120}$$

Il est important de noter que ces formules sont basées sur une approximation des parallèles en parabole mais des corrections supplémentaires sont prises en compte. Les longitudes sont considérées proportionnelles à la coordonnées UTM x.